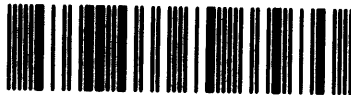


REFERENCE ONLY

SHL ITEM BARCODE



19 1783433 2

UNIVERSITY OF LONDON THESIS

Degree *PHD*

Year *2008*

Name of Author *CUNY, RACHA, DAWN.*

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

☐

This copy has been deposited in the Library of _____

☒

This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.

UNIVERSITY OF LONDON

SENATE HOUSE. MALET STREET, LONDON, WC1E 7HU



REPRODUCTION OF THESES

A thesis which is accepted by the University for the award of a Research Degree is placed in the Library of the College and in the University of London Library. The copyright of the thesis is retained by the author.

As you are about to submit a thesis for a Research Degree, you are required to sign the declaration below. This declaration is separate from any which may be made under arrangements with the College at which you have *pursued* your course (for internal candidates only). The declaration will be destroyed if your thesis is not approved by the examiners, being either rejected or referred for revision.

Academic Registrar

To be completed by the candidate

NAME IN FULL (please type surname in BLOCK CAPITALS)

RACHEL DAWN CURD

THESIS TITLE

MSP1 ANTIBODY SPECIFICITY AND MALARIA VACCINE DESIGN

DEGREE FOR WHICH THESIS IS PRESENTED Please select either... PhD

DATE OF AWARD OF DEGREE (To be completed by the University):

DECLARATION

1. I authorise that the thesis presented by me in *[2008] for examination for the MPhil/PhD Degree of the University of London shall, if a degree is awarded, be deposited in the library of the appropriate College and in the University of London Library and that, subject to the conditions set out below, my thesis be made available for public reference, inter-library loan and copying.
2. I authorise the College or University authorities as appropriate to supply a copy of the abstract of my thesis for inclusion in any published list of theses offered for higher degrees in British universities or in any supplement thereto, or for consultation in any central file of abstracts of such theses.
3. I authorise the College and the University of London Libraries, or their designated agents, to make a microform or digital copy of my thesis for the purposes of inter-library loan and the supply of copies.
4. I understand that before my thesis is made available for public reference, inter-library loan and copying, the following statement will have been included at the beginning of my thesis: The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.
5. I authorise the College and/or the University of London to make a microform or digital copy of my thesis in due course as the archival copy for permanent retention in substitution for the original copy.
6. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.
7. I understand that in the event of my thesis being not approved by the examiners, this declaration would become void.

***Please state year by hand, using a pen.**

DATE 17/10/07 SIGNATURE _____

Note: The University's Ordinances make provision for restriction of access to an MPhil/PhD thesis and/or the abstract but only in certain specified circumstances and for a maximum period of two years. If you wish to apply for such restriction, please enquire at your College about the conditions and procedures. External Students should enquire at the Research Degree Examinations Office, Room 261, Senate House.

MSP1 Antibody Specificity and Malaria Vaccine Design

Rachel Dawn Curd

UCL



UMI Number: U592541

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592541

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Rachel Dawn Curd, confirm that the work in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: _____

Abstract

This project focuses on a vaccine candidate MSP1₁₉ which is present in the asexual blood stages of the malaria parasite's lifecycle. Immunisation with MSP1₁₉ has been shown previously to protect against growth of the blood stage parasite. Previous studies have shown there are three types of antibodies produced against *Plasmodium falciparum* MSP1₁₉: inhibitory, blocking and neutral with only inhibitory antibodies giving protection. It is vital to identify the epitopes recognised by inhibitory antibodies to engineer an effective vaccine. The aim of this project is to map the antibody binding sites of *Plasmodium yoelii* (a rodent parasite) MSP1₁₉. Three protective *Plasmodium yoelii* MSP1₁₉ specific monoclonal antibodies that had been created previously were used. MSP1₁₉ variants containing amino acid changes in residues 12, 16, 17 and 28 were created and binding to the monoclonal antibodies was investigated using western blotting, ELISA and surface plasmon resonance analysis. This showed that all four residues were involved in antibody binding. A comparison of the residues found to be important for MSP1₁₉ antibody binding in *Plasmodium yoelii* and inhibitory antibody binding in *Plasmodium falciparum* show they lie within the same area. This suggests there are conserved areas for inhibitory antibody binding across the species implying a common mechanism of action. Immunisation studies with the MSP1₁₉ variants have shown that changes to residue 28 abolish the protective immune response to challenge infection with *Plasmodium yoelii* YM seen with wildtype MSP1₁₉. Structural NMR studies of wildtype and MSP1₁₉ variants have shown that residue 28 plays a vital structural role. The information presented in this project could be important in developing antigens for vaccination to specifically stimulate production of inhibitory antibodies. It could help direct research into understanding the mechanism of action of inhibitory antibodies and aid in the development of new therapeutic strategies targeting MSP1₁₉.

Table of Contents

Abstract	3
List of Figures	9
Abbreviations.....	15
Chapter 1: Introduction.....	19
1.1 Malaria.....	19
1.2 Current and future malaria treatments	20
1.3 Vaccines	23
1.3.1 Vaccine Candidates	24
1.4 Introduction to MSP1 ₁₉	28
1.5 Understanding the structure of the protein	29
1.5.1 Structure of MSP1 ₁₉	30
1.6 Immunity to MSP1 ₁₉	32
1.6.1 Evidence from animal studies of MSP1 ₁₉ as a vaccine candidate.....	32
1.6.2 Immunity to MSP1 ₁₉ in humans	33
1.6.3 Studies on <i>P. falciparum</i> MSP1 ₁₉ antibody binding	35
1.6.4 Studies on <i>P. yoelii</i> MSP1 ₁₉ antibodies and antibody binding.....	36
1.7 Aims of the project	42
Chapter 2: Materials and Methods.....	43
2.1 Materials	43
2.1.1 Buffers and solutions.....	43
2.1.2 Bacterial cell culture: Bacteria and Plasmids	44
2.2 Methods: Binding studies.....	45
2.2.1 Production of MSP1 ₁₉ variant clones as GST fusion proteins	45
2.2.2 Expression of GST-MSP1 ₁₉ variants.....	46
2.2.3 Purification of GST-MSP1 ₁₉ variants.....	47
2.2.4 Quantification of GST-MSP1 ₁₉ variants	47
2.2.5 Western blotting	48
2.2.5 ELISA analysis of GST-MSP1 ₁₉ variant proteins	49
2.2.6 Surface plasmon resonance (SPR) analysis.....	49
2.2.7 Molecular modelling of <i>P. yoelii</i> MSP1 ₁₉	50
2.2.8 <i>In silico</i> variation of residues in <i>P. yoelii</i> MSP1 ₁₉	50
2.3 Methods: Immunisation Studies	51

2.3.1	ELISA analysis of serum samples from immunisation studies	52
2.4	Methods: Protein preparation for structural NMR studies	53
2.4.1	Production of MSP1 ₁₉ variant clones as his-tagged proteins	53
2.4.2	Production of recodonised MSP1 ₁₉ variant genes	53
2.4.3	Preparation of DNA for electroporation into <i>Pichia pastoris</i>	53
2.4.4	Restriction enzyme digestion	53
2.4.5	Dephosphorylation of DNA	54
2.4.6	Agarose gel electrophoresis.....	54
2.4.7	Agarose gel extraction.....	55
2.4.8	Ligation reactions	55
2.4.9	Transformation of chemically competent <i>E. coli</i>	55
2.4.10	Production of <i>E. coli</i> glycerol stocks.....	56
2.4.11	Large scale DNA purification	56
2.4.12	Production of Glu28→Gln his-MSP1 ₁₉ variant using Quikchange XL site-directed mutagenesis kit	56
2.4.13	Linerisation of pPIC9K for transformation into <i>Pichia pastoris</i>	57
2.4.14	Preparation of electrocompetent <i>Pichia pastoris</i>	57
2.4.15	Geneticin screening of transformants.....	58
2.4.16	Small scale expression time course in <i>P. pastoris</i>	59
2.4.17	Large scale expression of his-tagged MSP1 ₁₉ variants.....	60
2.4.18	Large scale purification of his-tagged MSP1 ₁₉ variants	61
2.4.19	Quantification of his-MSP1 ₁₉ variants	62
2.5	Methods: Nuclear Magnetic Resonance (NMR) Spectroscopy.....	67
2.5.1	Preparation of his-MSP1 ₁₉ variants for NMR spectroscopy	67
2.5.2	¹⁵ N Heteronuclear Single Quantum Correlation (¹⁵ N-HSQC) Spectroscopy	67
2.5.3	Using NMR to determine the 3D structure of wildtype his-MSP1 ₁₉ and Glu28→Lys his-MSP1 ₁₉	68
2.5.4	Structure determination using ARIA 1.2.....	69
Chapter 3: Antibody binding studies on individual amino acid variants		72
3.1	Introduction	72
3.2	Expression and purification of GST-MSP1 ₁₉ variants	77
3.3	Western blotting analysis of antibody binding to GST-MSP1 ₁₉ variants	77

3.4	ELISA analysis of antibody binding to GST-MSP1 ₁₉ variants.....	78
3.5	SPR of antibody binding to GST-MSP1 ₁₉ variants	79
3.6	Discussion	103
Chapter 4: Designing a double MSP1₁₉ variant to affect all three monoclonal antibodies		105
4.1	Introduction	105
4.2	Expression and purification of double Lys16→Glu /Glu28→Lys GST-MSP1 ₁₉ variant	105
4.3	Western blotting analysis of antibody binding to residues 16 and 28 single and double GST-MSP1 ₁₉ variants	106
4.4	ELISA analysis of antibody binding to residues 16 and 28 double and single GST-MSP1 ₁₉ variants	107
4.5	Discussion	118
Chapter 5: Immunisation studies 1 – do the MSP1₁₉ variations affect protection?		119
5.1	Introduction	119
5.2	Immunisation studies with MSP1 ₁₉ variants	121
5.3	ELISA analysis of antibody titres following immunisation with MSP1 ₁₉ variants	122
5.4	Discussion	135
Chapter 6: Structural analysis of MSP1₁₉ variants.....		138
6.1	Introduction	138
6.2	Molecular modelling of <i>P. yoelii</i> MSP1 ₁₉	138
6.3	Expression and purification of his-MSP1 ₁₉ variants	140
6.4	¹⁵ N-HSQC NMR analysis of his-MSP1 ₁₉ variants	142
6.5	Discussion	174
Chapter 7: Designing and analysing a Glu28→Gln MSP1₁₉ variant.....		177
7.1	Introduction	177
7.2	Expression and purification of Glu28→Gln GST-MSP1 ₁₉ variant.....	179
7.3	Western blotting analysis of antibody binding to residue 28 GST-MSP1 ₁₉ variants	179
7.4	ELISA analysis of antibody binding to residue 28 GST-MSP1 ₁₉ variants ...	180
7.5	Expression and purification of Glu28→Gln his-MSP1 ₁₉ variant.....	181

7.6	¹⁵ N-HSQC NMR analysis of Glu28→Gln his-MSP1 ₁₉ variant	181
7.7	Discussion	201
Chapter 8: Further immunisation studies 2 – do the residue 28 variants affect protection differently?		203
8.1	Introduction	203
8.2	Immunisation studies with residue 28 MSP1 ₁₉ variants.....	204
8.3	ELISA analysis of antibody titres following immunisation with residue 28 MSP1 ₁₉ variants.....	205
8.4	ELISA analysis of serum antibodies against wildtype MSP1 ₁₉ and residue 28 MSP1 ₁₉ variants	207
8.5	Discussion	228
Chapter 9: Solving the 3D structures of wildtype <i>P. yoelii</i> MSP1₁₉ and Glu28→Lys MSP1₁₉ variant using NMR.....		231
9.1	Introduction	231
9.2	Expression and Purification of ¹³ C/ ¹⁵ N labelled his-MSP1 ₁₉	232
9.3	Assigning NMR Spectra for wildtype and Glu28→Lys MSP1 ₁₉ variant.....	233
9.3.1	HNCACB and CBCACONH NMR spectra	233
9.3.2	HCCCONH and HCCH-TOCSY NMR spectra.....	234
9.3.3	Determining distance restraints for wildtype and Glu28→Lys MSP1 ₁₉	235
9.3.4	Predicting Phi and Psi angles using TALOS.....	236
9.3.5	D ₂ O exchange analysis.....	237
9.4	Calculating the 3D structure of wildtype and Glu28→Lys MSP1 ₁₉ variant	239
9.4.1	ARIA	239
9.4.2	Evaluating the quality of the NMR structures	240
9.5	Comparing the 3D NMR structure of wildtype <i>P. yoelii</i> MSP1 ₁₉ to the homology model.....	241
9.6	Comparing the 3D NMR structure of wildtype and Glu28→Lys MSP1 ₁₉ variant	241
9.7	Discussion	285
Chapter 10: Discussion		288
10.1	Introduction	288
10.2	Electrostatic potential of wildtype and Glu28→Lys MSP1 ₁₉ variant	290

10.3	Comparison of electrostatic potential of wildtype MSP1 ₁₉ from different species.....	290
10.4	Discussion	301
10.4.1	Structure function relationship of wildtype <i>P. yoelii</i> MSP1 ₁₉ and Glu28→Lys MSP1 ₁₉ variant	301
10.4.2	Structural comparison between <i>P. falciparum</i> and <i>P. yoelii</i>	302
10.4.3	Comparison of <i>P. yoelii</i> and <i>P. falciparum</i> antibody binding sites.....	303
10.5	Overall implications of this work	309
10.6	Future Work	311
Acknowledgements		313
Bibliography.....		314

List of Figures

Figure 1.1: Current and future interventions for malaria control.	22
Figure 1.2: Life cycle of <i>P. falciparum</i> highlighting the immune mechanisms for different types of vaccines.	26
Figure 1.3: Schematic showing primary and secondary processing of MSP1.	38
Figure 1.4: Comparison of the backbone of MSP1 ₁₉ from different <i>plasmodium</i> species.	39
Figure 1.5: Alignment of <i>P. yoelii</i> (Py) and <i>P. falciparum</i> (Pf) MSP1 ₁₉ highlighting the conserved residues.	41
Figure 2.1: his-MSP1 ₁₉ variants recodonised gene sequence.....	63
Figure 2.2: Overview of the strategy used to prepare the pPIC9K and his-MSP1 ₁₉ recodonised genes for transformation into yeast.....	65
Figure 3.1: Overview of the site-directed mutagenesis protocol used to produce the amino acid variations in MSP1 ₁₉	75
Figure 3.2: Single amino acid variations made to wildtype MSP1 ₁₉	81
Figure 3.3: NuPAGE gel analysis of Asn17→His GST-MSP1 ₁₉ variant expression and purification.	83
Figure 3.4: NuPAGE gel analysis of 500 ng of GST-MSP1 ₁₉ variants quantified by densitometry.	84
Figure 3.5: Western blotting analysis of antibody binding to GST-MSP1 ₁₉ variants.	85
Figure 3.6: Antibody sandwich ELISA to determine the optimum concentration of capture antibody.	87
Figure 3.7: ELISA of B6 antibody binding to GST-MSP1 ₁₉ variants.....	88
Figure 3.8: ELISA of F5 antibody binding to GST-MSP1 ₁₉ variants.	90
Figure 3.9: ELISA of B10 antibody binding to GST-MSP1 ₁₉ variants.	92
Figure 3.10: Overview of surface plasmon resonance.	94
Figure 3.11: Schematic of surface plasmon resonance sensorgram obtained in the experiments in this project.	96
Figure 3.12: SPR of B6 binding to GST-MSP1 ₁₉ variants.	98
Figure 3.13: SPR of F5 binding to GST-MSP1 ₁₉ variants.....	99
Figure 3.14: SPR of B10 binding to GST-MSP1 ₁₉ variants.	100

Figure 3.15: Summary of the effect of single amino acid changes on B6, F5 and B10 antibody binding to GST-MSP1 ₁₉	101
Figure 4.1: NuPAGE gel analysis of double Lys16→Glu/Glu28→Lys GST-MSP1 ₁₉ variant expression and purification.	108
Figure 4.2: Western blotting analysis of antibody binding to residues 16 and 28 single and double GST-MSP1 ₁₉ variants.	110
Figure 4.3: ELISA of B6 antibody binding to residues 16 and 28 single and double GST-MSP1 ₁₉ variants.	112
Figure 4.4: ELISA of F5 antibody binding to residues 16 and 28 single and double GST-MSP1 ₁₉ variants.	114
Figure 4.5: ELISA of B10 antibody binding to residues 16 and 28 single and double GST-MSP1 ₁₉ variants.	116
Figure 5.1: Course of <i>P. yoelii</i> YM infection in mice immunised with wildtype MSP1 ₁₉	124
Figure 5.2: Course of <i>P. yoelii</i> YM infection in mice immunised with purified GST..	125
Figure 5.3: Course of <i>P. yoelii</i> YM infection in mice immunised with Arg 12→Leu MSP1 ₁₉ variant.....	126
Figure 5.4: Course of <i>P. yoelii</i> YM infection in mice immunised with Asn17→His MSP1 ₁₉ variant.....	127
Figure 5.5: Course of <i>P. yoelii</i> YM infection in mice immunised with Lys16→Glu MSP1 ₁₉ variant.....	128
Figure 5.6: Course of <i>P. yoelii</i> YM infection in mice immunised with Glu28→Lys MSP1 ₁₉ variant.....	129
Figure 5.7: Course of <i>P. yoelii</i> YM infection in mice immunised with double Lys16→Glu/Glu28→Lys MSP1 ₁₉ variant.....	130
Figure 5.8: Course of <i>P. yoelii</i> YM infection in groups of mice immunised with wildtype and MSP1 ₁₉ variants.	131
Figure 5.9: Antibody binding curves for pooled serum from groups of mice immunised with wildtype and MSP1 ₁₉ variants against his-tagged wildtype MSP1 ₁₉	133
Figure 6.1: <i>In silico</i> variation of residues in <i>P. yoelii</i> MSP1 ₁₉ model.	146
Figure 6.2: NuPAGE gel analysis of the expression time course for his-MSP1 ₁₉ variants in <i>Pichia pastoris</i>	148

Figure 6.3: NuPAGE gel analysis of the purification of 50 ml cultures of wildtype and Glu28→Lys MSP1 ₁₉ variant his-tagged proteins expressed in <i>Pichia pastoris</i>	150
Figure 6.4: NuPAGE gel analysis of the purification of ¹⁵ N labelled wildtype and Glu28→Lys MSP1 ₁₉ variant his-tagged proteins for NMR analysis.	152
Figure 6.5: 1D ¹ H NMR Spectra for wildtype MSP1 ₁₉ and Glu28→Lys MSP1 ₁₉ proteins.	154
Figure 6.6: ¹⁵ N-HSQC Spectrum of wildtype MSP1 ₁₉	156
Figure 6.7: Comparison of Arg12→Leu MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	158
Figure 6.8: Comparison of Lys16→Glu MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	160
Figure 6.9: Comparison of Asn17→His MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	162
Figure 6.10: Comparison of Glu28→Lys MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	164
Figure 6.11: Comparison of double Lys16→Glu/Glu28→Lys MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.....	166
Figure 6.12: Comparison of double Lys16→Glu/Glu28→Lys and Glu28→Lys MSP1 ₁₉ variants ¹⁵ N-HSQC NMR spectra.	168
Figure 6.13: Comparison of the combined ¹⁵ N and ¹ H chemical shift differences of double Lys16→Glu/Glu28→Lys, Glu28→Lys and Lys16→Glu MSP1 ₁₉ variants compared to wildtype MSP1 ₁₉	170
Figure 6.14: Mapping the NH peaks that have moved in the ¹⁵ N-HSQC spectra onto the 3D structure of <i>P. yoelii</i> MSP1 ₁₉	172
Figure 7.1: Differences between glutamic acid, lysine and glutamine side chains.....	178
Figure 7.2: Western blotting analysis of antibody binding to residue 28 GST-MSP1 ₁₉ variants.	183
Figure 7.3: ELISA of B6 antibody binding to residue 28 GST-MSP1 ₁₉ variants.	185
Figure 7.4: ELISA of F5 antibody binding to residue 28 GST-MSP1 ₁₉ variants.....	187
Figure 7.5: ELISA of B10 antibody binding to residue 28 GST-MSP1 ₁₉ variants.	189
Figure 7.6: Comparison of Glu28→Gln MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	191

Figure 7.7: Mapping the NH peaks that have moved in the Glu28→Gln MSP1 ₁₉ variant ¹⁵ N-HSQC spectra onto the 3D structure of <i>P. yoelii</i> MSP1 ₁₉	193
Figure 7.8: Comparison of Glu28→Gln MSP1 ₁₉ variant, Glu28→Lys MSP1 ₁₉ variant and wildtype MSP1 ₁₉ ¹⁵ N-HSQC NMR spectra.	195
Figure 7.9: Comparison of the combined ¹⁵ N and ¹ H chemical shift differences of Glu28→Gln and Glu28→Lys MSP1 ₁₉ variants compared to wildtype MSP1 ₁₉	197
Figure 7.10: Differences between the combined ¹⁵ N and ¹ H chemical shift differences of Glu28→Gln and Glu28→Lys MSP1 ₁₉ variants compared to wildtype MSP1 ₁₉	199
Figure 8.1: Course of <i>P. yoelii</i> infection in mice immunised with wildtype MSP1 ₁₉ and antibody binding curves.	209
Figure 8.2: Course of <i>P. yoelii</i> infection in mice immunised with GST.	211
Figure 8.3: Course of <i>P. yoelii</i> infection in mice immunised with Glu28→Lys MSP1 ₁₉ variant and antibody binding curves.	213
Figure 8.4: Course of <i>P. yoelii</i> infection in mice immunised with the double Lys16→Glu/Glu28→Lys MSP1 ₁₉ variant and antibody binding curves.	215
Figure 8.5: Course of <i>P. yoelii</i> infection in mice immunised with Glu28→Gln MSP1 ₁₉ variant and antibody binding curves.	217
Figure 8.6: Course of <i>P. yoelii</i> YM infection in groups of mice immunised with wildtype and residue 28 MSP1 ₁₉ variants and antibody binding curves.	219
Figure 8.7: Antibody binding curves for serum from mice immunised with Glu28→Lys MSP1 ₁₉ variant against his-tagged wildtype and Glu28→Lys MSP1 ₁₉ variant.	221
Figure 8.8: Antibody binding curves for serum from mice immunised with double Lys16→Glu/Glu28→Lys MSP1 ₁₉ variant against his-tagged wildtype and double Lys16→Glu/Glu28→Lys MSP1 ₁₉ variant.	223
Figure 8.9: Antibody binding curves for serum from mice immunised with Glu28→Gln MSP1 ₁₉ variant against his-tagged wildtype and Glu28→Gln MSP1 ₁₉ variant.	225
Figure 8.10: Antibody binding curves for serum from mice immunised with wildtype MSP1 ₁₉ against his-tagged residue 28 MSP1 ₁₉ variants.	227
Figure 9.1: NuPAGE gel analysis of expression of doubly labelled ¹³ C/ ¹⁵ N MSP1 ₁₉ proteins in <i>P. pastoris</i>	244
Figure 9.2: Schematic representation of how NMR spectroscopy was used for structural determination of wildtype MSP1 ₁₉ and Glu28→Lys MSP1 ₁₉ variants.	246
Figure 9.3: HNCACB and CBCACONH spectra.	248

Figure 9.4: Assignment of residues 55 to 65 in the HNCACB spectrum of wildtype MSP1 ₁₉	250
Figure 9.5: Assigned ¹⁵ N-HSQC spectra for wildtype and Glu28→Lys MSP1 ₁₉ variant.	252
Figure 9.6: Assignment of side chain ¹ H of glycine in HCCCONH spectrum.	254
Figure 9.7: Assignment of side chain ¹ H of asparagine in HCCCONH spectrum.	255
Figure 9.8: Assignment of side chain ¹ H of isoleucine in HCCCONH spectrum.	256
Figure 9.9: HCCH-TOCSY NMR spectrum.	257
Figure 9.10: Schematic representation of a 3D ¹⁵ N-HSQC-NOESY experiment.	259
Figure 9.11: ¹⁵ N-HSQC NOESY spectrum for wildtype MSP1 ₁₉ residues 20 to 22 highlighting the NOEs between residues in a β-sheet.	260
Figure 9.12: ¹⁵ N-HSQC NOESY spectrum for wildtype MSP1 ₁₉ residues 9 to 11 highlighting the NOEs between residues involved in a turn.	262
Figure 9.13: Histogram of D ₂ O exchange rates for wildtype and Glu28→Lys MSP1 ₁₉ variant.	264
Figure 9.14: The residues protected from D ₂ O exchange mapped onto the best energy wildtype MSP1 ₁₉ NMR structure.	266
Figure 9.15: Schematic overview of the operation performed by ARIA.	268
Figure 9.16: 20 best energy structures of wildtype and Glu28→Lys MSP1 ₁₉ variant.	269
Figure 9.17: Ramachandran plot of the 20 best energy wildtype MSP1 ₁₉ structures.	271
Figure 9.18: Ramachandran plot of the best energy wildtype MSP1 ₁₉ structure.	272
Figure 9.19: Ramachandran plot of the 20 best energy Glu28→Lys MSP1 ₁₉ variant structures.	273
Figure 9.20: Ramachandran plot of the best energy Glu28→Lys MSP1 ₁₉ variant structure.	274
Figure 9.21: Comparison of the best energy wildtype MSP1 ₁₉ NMR structure and the homology model.	275
Figure 9.22: Comparison between the backbone structure of the first EGF domain of wildtype and Glu28→Lys MSP1 ₁₉	277
Figure 9.23: Comparison of the orientation of the residues in wildtype and Glu28→Lys MSP1 ₁₉ variant that have moved 0.2 ppm in the ¹⁵ N-HSQC spectrum.	279
Figure 9.24: Comparison of the orientation of the charged residues in wildtype and Glu28→Lys MSP1 ₁₉ variant that have moved 0.2 ppm in the ¹⁵ N-HSQC spectrum.	281

Figure 9.25: Comparison of the orientation of the aromatic residues in wildtype and Glu28→Lys MSP1 ₁₉ variant.	283
Figure 10.1: Comparison of electrostatic potential of wildtype and Glu28→Lys MSP1 ₁₉ variant.....	292
Figure 10.2: Alignment of <i>P. yoelii</i> (Py) and <i>P. falciparum</i> (Pf) MSP1 ₁₉ highlighting the conserved residues.....	294
Figure 10.3 Comparison of <i>P. yoelii</i> and <i>P. falciparum</i> MSP1 ₁₉ NMR structures.	295
Figure 10.4: Comparison of electrostatic potential of <i>P. yoelii</i> and <i>P. falciparum</i> MSP1 ₁₉	297
Figure 10.5: Comparison of electrostatic potential of MSP1 ₁₉ from different species..	298
Figure 10.6: Alignment of <i>P. yoelii</i> (Py) and <i>P. falciparum</i> (Pf) MSP1 ₁₉ highlighting the four <i>P. yoelii</i> variants produced in this thesis relative to residues shown to affect <i>P. falciparum</i> MSP1 ₁₉ inhibitory antibody binding.	306
Figure 10.7: Important residues for inhibitory antibody binding in <i>P. falciparum</i> MSP1 ₁₉ mapped onto the NMR structure of <i>P. yoelii</i> MSP1 ₁₉	307

Abbreviations

1D	1 Dimensional
¹ H	1 Hydrogen
¹³ C	13 Carbon
¹⁵ N	15 Nitrogen
¹⁵ N-HSQC	¹⁵ N Heteronuclear Single Quantum Correlation Spectroscopy
MSP1 ₁₉	19 KDa processing fragment of MSP1
2D	2 Dimensional
3D	3 Dimensional
MOPS	3-(N-morpholino)propanesulphuric acid
AOX1	alcohol oxidase gene
ARIA	Ambiguous Restraint for Iterative Assignment
NH	amide proton
Å	angstroms
BSA	Bovine serum albumin
BMGlc	Buffered Glucose-complex Medium
BMMY	Buffered Methanol-complex Medium
CIP	Calf intestinal alkaline phosphatase
Arg12→Leu	change from arginine to leucine at residue 12
Asn17→His	change from asparagine to histidine at residue 17
Glu28→Gln	change from glutamic acid to glutamine at residue 28
Glu28→Ile	change from glutamic acid to isoleucine at residue 28
Glu28→Lys	change from glutamic acid to lysine at residue 28
Lys16→Glu	change from lysine to glutamic acid at

Da	residue 16
DNA	Daltons
$^2\text{H}_2\text{O}$	deoxyribonucleic acid
D_2O	deuterium oxide
dH_2O	deuterium oxide
DTT	distilled water
ELISA	dithiothreitol
<i>E. coli</i>	Enzyme-Linked Immunosorbent Assay
EDTA	<i>Escherichia coli</i>
FCA	ethylenediaminetetraacetic acid
FIA	Freund's complete adjuvant
GST	Freund's incomplete adjuvant
GPI	Glutathione S-transferase
GST-MSP1 ₁₉	glycophosphatidylinositol
his-MSP1 ₁₉	GST-fusion MSP1 ₁₉ protein
HRP	hexahistidine tagged MSP1 ₁₉
HCl	Horse radish peroxidase
H-bond	Hydrochloric acid
IgG	Hydrogen bond
i.p.	immunoglobulins
i.v	intraperiton
IPTG	intravenous
kDa	Isopropyl-beta-D-thiogalactopyranoside
kV	kilodaltons
LB	kilovolts
MHz	Luria-Bertani Broth
MSP	Mega Hertz
μg	Merozoite surface protein
μl	Microgrammes
mA	Microlitres
mg	milliamps
ml	Milligrammes
	Millilitres

mM	Millimolar
ms	milliseconds
min	Minute
MWCO	Molecular weight cut off
MOLMOL	Molecule analysis and Molecular display
ng	nanograms
nm	nanometres
Ni-NTA	nickel-nitrilotriacetic acid
NP40	Nonidet P40
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
Ω	ohms
OD	Optical Density
ppm	parts per million
%	Percent
PBS	Phosphate buffered saline
<i>P. pastoris</i>	<i>Pichia pastoris</i>
<i>P. falciparum</i>	Plasmodium falciparum
<i>P. yoelii</i>	Plasmodium yoelii
KCl	potassium chloride
PBD	Protein Data Bank
g	relative centrifugal force
rpm	revolutions per minute
SPR	Surface plasmon resonance
TALOS	Torsion Angle Likelihood Obtained from Shift and Sequence similarity
TOCSY	Total correlation spectroscopy
TROSY	Transverse Relaxation Optimised Spectroscopy
TAE	Tris-acetate
v/v	Volume for volume
WHO	World Health Organisation

w/v	Weight for volume
YEPD	Yeast Extract Peptone Dextrose Medium
YNB	Yeast Nitrogen Base
YND	Yeast Nitrogen Dextrose

Chapter 1: Introduction

1.1 Malaria

Malaria is a devastating disease. The World Health Organisation (WHO) estimates that it causes up to 3 million deaths annually and 2.4 billion people, namely half the world's population is at risk (APPMG, 2005). A large proportion of the malaria deaths are of children in Africa where malaria-attributable death rates for children under 5 have been reported as high as 25 to 30 % of the total. Malaria is also a large economic burden with countries with intensive malaria lagging in growth by 1.3 % per person compared to neighbouring non-endemic countries (Breman *et al.*, 2001).

Malaria is caused by a protozoan called *Plasmodium* and is transmitted by mosquitoes. There are four species of *Plasmodium* that infect humans: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. *P. falciparum* is the most virulent species that affects humans. There is a wide range of symptoms associated with malaria, and the cause of these symptoms is not always obvious, which can lead to misdiagnosis. If a patient with malaria is not treated the patient can deteriorate rapidly as the parasite rapidly replicates in the blood stream during the infection (Breman, 2001). The symptoms tend to occur in cycles of fever and chills which correspond with phases of the parasite's life cycle. Malaria can lead onto severe malaria and this can include: shock, respiratory distress, anaemia, hypoglycaemia and cerebral malaria. The severe malaria can lead to death. Malaria can also have long term effects leading to impaired growth and development and increased infant mortality. Malaria can also cause significant problems for pregnant mothers with the mother suffering from anaemia and hypoglycaemia and the baby being born with a low birth weight and increased risk of infant mortality (Breman, 2001, Chiang *et al.*, 2006).

1.2 Current and future malaria treatments

Figure 1.1 highlights the current and future directions of malaria treatments. The main focus of malaria prevention in the early 1960s was to use insecticides such as DDT to kill the mosquitoes. The mosquitoes developed resistance to the insecticides and this approach was abandoned by the WHO in 1969. Another approach that is being used to prevent malaria infection is using insecticide impregnated bednets to stop mosquitoes biting people at night and transmitting the disease. This approach is very effective but relies on ensuring the bednets are used correctly and are not damaged. Anti-malarial drugs have long been used for the treatment of malaria and are still being used.

Examples of anti-malaria drugs include: Chloroquine, Quinine and Mefloquine. The main problem with the anti-malarial drugs is that the *Plasmodium* parasite is developing resistance to them. For example, *P. falciparum* has developed resistance to Chloroquine but it is still widely prescribed because it is cheaper than other anti-malarials and is well tolerated. A newer anti-malarial drug, Artemisinin has been developed which is able to work against multi-drug resistant *P. falciparum* but is more expensive than the older drugs (Bremar, 2001, Chiang *et al.*, 2006). In order for anti-malarial drugs to be used in the future new drugs that work by different mechanisms need to be developed that will overcome the problem of drug resistance. Any new anti-malaria drugs that are produced need to be affordable for malaria endemic countries.

Potential future interventions for malaria control and treatment include: the genetic modification of vectors and the development of vaccines. The genetic modification of vectors would involve genetically modifying mosquitoes so that they were unable to support the lifecycle of the *Plasmodium* parasite or so that the parasite would die if the mosquito ingested it in the blood meal. The aim would be to prevent malaria transmission. The genetic modification of mosquitoes may have environmental consequences and be unpopular. Work is being carried out to develop vaccines to prevent transmission of malaria, develop immunity to malaria and to reduce the severity of symptoms (Bremar, 2001).

In addition to new treatments malaria control could involve better education and social and political awareness. For example, modifying attitudes to the environment during engineering projects such as building dams and new residential areas where this could increase mosquito breeding grounds and bring more people into close contact with mosquitoes. Altering people's behaviour and attitude towards malaria in endemic countries could also help in the future including improving education about the condition, draining swamps and eliminating other mosquito breeding grounds and improving availability of rapid diagnostic tests for malaria (Breman, 2001).

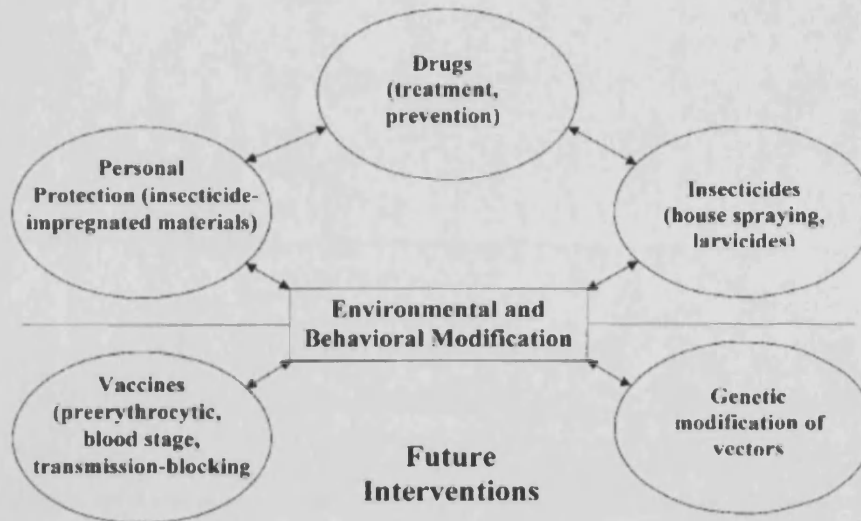


Figure 1.1: Current and future interventions for malaria control.

Figure from Breman *et al.* (Breman, 2001)

1.3 Vaccines

Vaccines could provide a cost effective method for the control and prevention of malaria. Vaccines could be targeted at a number of different groups of people to help in controlling the disease including: pregnant women and women of child-bearing age to prevent malaria during pregnancy; babies and children in malaria endemic countries that have not yet built up any natural immunity to malaria infection and visitors to malaria endemic countries.

There are a number of arguments that have led scientists to believe that the development of an effective malaria vaccine will be possible. Protective immunity to malaria has been induced in every animal model of the disease. In humans, the level of parasitaemia and the severity of malaria decrease with age suggesting that repeated infection can induce immune responses that affect the parasitaemia and severity of the disease. Understanding the way in which the repeated infections induce a protective immune response could lead to the development of a vaccine to mimic this induction of immunity. Studies have shown that protection against *P. falciparum* can be passively transferred in humans by immunisation with immunoglobulin purified from the blood of adults that have been living in malaria endemic countries for their whole lives. The studies have shown that antibodies against the blood stage antigens can protect against *P. falciparum* parasite growth in the blood stream and this could mean that if the antigens that produce the immune response could be identified and purified for vaccination they could potentially induce the same protective immune response (Hoffman & Miller, 1996). The final reason to believe that a protective vaccine would be possible is from immunisation studies with radiation-attenuated sporozoites in mice and humans. In these studies it was found that radiation-attenuated sporozoites induce sterile protective immunity. This could suggest that vaccines based on these parasites or the proteins they contain could help protect against the disease (Hoffman & Miller, 1996).

1.3.1 Vaccine Candidates

The lifecycle of the *Plasmodium* parasite is made up of a number of distinct stages, vaccines are therefore being developed to target the different stages of the lifecycle as highlighted in figure 1.2. Three main classes of vaccines are being developed: pre-erythrocytic vaccines, erythrocytic vaccines and transmission blocking vaccines.

The aim of developing a pre-erythrocytic vaccine is to stop the parasite when it enters the body before it has a chance to get to the erythrocytes. Examples of antigens for this type of vaccine include circumsporozoite protein, liver stage antigen-1 and liver stage antigen-3. Another approach that is being taken to develop a pre-erythrocytic vaccine is by a company called Sanaria who are working on producing irradiated sporozoites for use as a vaccine (Chiang *et al.*, 2006).

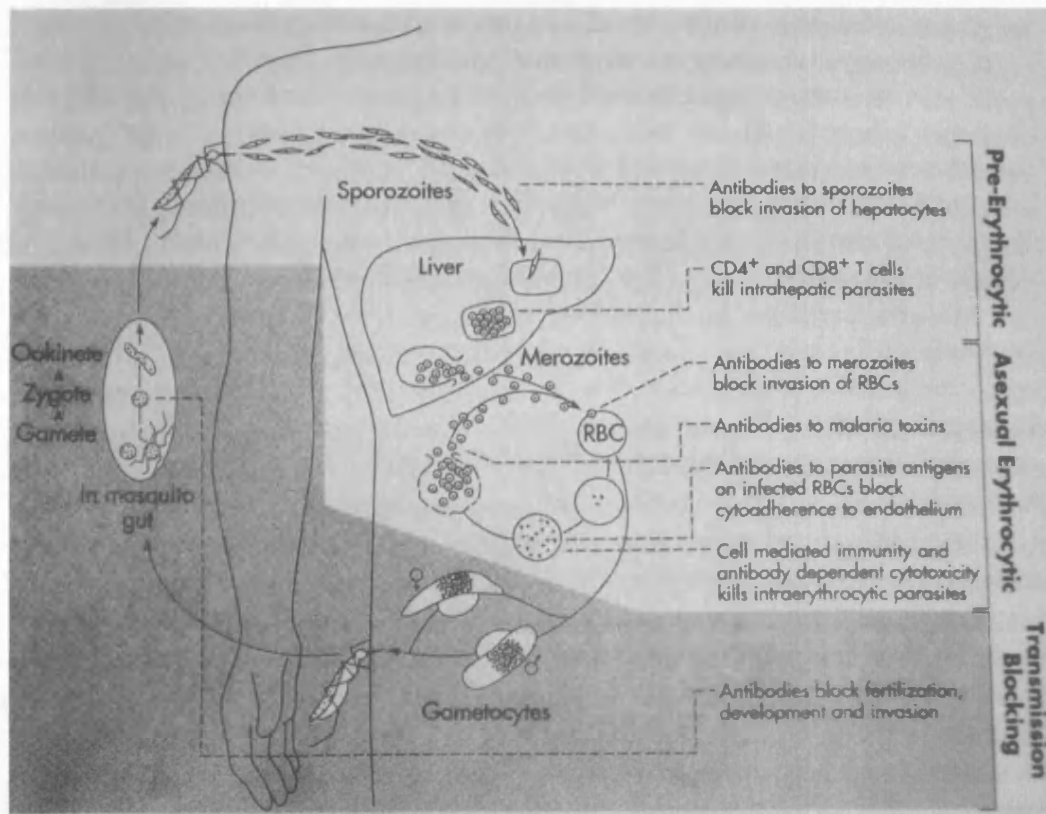
The aim of an erythrocytic vaccine is to target the asexual erythrocytic parasites. The asexual erythrocytic parasites cause the clinical symptoms of malaria and the severity of the disease is related to the parasitaemia. An erythrocytic vaccine could reduce the severity of the malaria infection, for example, by inducing an antibody response that targets proteins on the surface of the merozoite thus blocking erythrocyte invasion and stopping the replication of the merozoite in the erythrocyte. The advantage of this type of vaccine is it would not be required to induce complete resistance to infection to be effective (Miller *et al.*, 1986). Examples of antigens being developed for this type of vaccine include: recombinant proteins based on fragments of merozoite surface protein-1, AMA-1 and merozoite surface protein-3.

The aim of transmission blocking vaccines is to induce the production of antibodies in the human host that will be ingested in the mosquito's blood meal. The antibodies will react with newly expressed proteins on the surface of the parasite in the mosquito gut and prevent transmission of the parasite to the next person the mosquito bites. The vaccine may do this by inducing antibodies that stop the migration of the parasite across the mosquito midgut. The advantage of this type of vaccine is that the antigens may be conserved as the sexual stage parasites do not need to evade the immune system in the human host, but the main disadvantage is that vaccines will not protect the individual

from getting malaria (Chiang *et al.*, 2006). In addition to the vaccines aimed at the individual stages of the parasite life-cycle, multi-protein vaccines are also being developed that combine proteins from different lifecycle stages (Hoffman & Miller, 1996). Novel types of vaccines are also being developed that do not rely on recombinant protein production (which can be difficult with malaria parasite proteins) including DNA based vaccines, genetically attenuated parasites and edible plant vaccines. There is currently no effective malaria vaccine on the market and producing an effective vaccine remains a huge future challenge for the malaria vaccine community.

Figure 1.2: Life cycle of *P. falciparum* highlighting the immune mechanisms for different types of vaccines.

Figure from Hoffman *et al.* (Hoffman & Miller, 1996)



1.4 Introduction to MSP1₁₉

This project focuses on a leading vaccine candidate that is present in the asexual blood stage of the malaria parasite's life cycle. This is the stage that causes the clinical symptoms. The protein is a 19 kDa C-terminal region of the major merozoite surface protein 1 (MSP1), designated MSP1₁₉. MSP1 undergoes two processing events as shown in figure 1.3. The exact function of MSP1 and the processing events are still unclear. It has been suggested that MSP1 is involved in binding to the surface of erythrocytes and is involved in recognition of and attachment to these cells (Holder & Blackman, 1994). The primary processing event occurs at the end of schizogony around the time of merozoite release. Primary processing produces four polypeptides MSP1₈₃, MSP1₃₀, MSP1₃₈ and MSP1₄₂ (Holder *et al.*, 1994). The MSP1₄₂ portion of the protein is anchored to the membrane by a glycosyl phosphatidyl inositol (GPI) moiety and polypeptides are held in a non-covalent complex on the merozoite surface (McBride & Heidrich, 1987). The MSP1 complex undergoes a secondary processing event at or immediately before erythrocyte invasion where MSP1₄₂ polypeptide is cleaved by proteases to MSP1₃₃ and MSP1₁₉. The MSP1₃₃ polypeptide is shed from the surface of the parasite along with MSP1₈₃, MSP1₃₀ and MSP1₃₈. MSP1₁₉ is carried on the surface of the invading parasite into the erythrocyte (Blackman & Holder, 1992, Blackman *et al.*, 1991). The exact function of this MSP1₁₉ fragment is unclear. Hypotheses that have been put forward are that MSP1₁₉ could interact with receptor proteins or could provide down stream biochemical communication in the cells. It has been suggested that MSP1₁₉ could be involved in a signalling cascade to indicate that the cell has been invaded by the parasite and therefore switch on red blood cell machinery (Holder & Blackman, 1994).

It is clear that the secondary processing event needs to occur in order for the merozoite to invade the erythrocyte. It has been shown by Blackman *et al.* that antibodies to MSP1₁₉ inhibit erythrocyte invasion and that MSP1₁₉ can induce an immune response that protects against growth of the blood stage parasite (Blackman *et al.*, 1990, Blackman *et al.*, 1994).

In order to understand and utilise MSP1₁₉ as a protein for vaccine development you must understand the structure of the protein and immunity to that protein.

1.5 Understanding the structure of the protein

Understanding the structure of MSP1₁₉ is vital for its use in developing a vaccine or in rational drug design. Analysis of the three-dimensional structure of the protein can help to understand how the biological function of the protein is related to the structure and can help identify ligand, receptor and antibody binding sites. Understanding the three-dimensional structure of MSP1₁₉ for vaccine design is particularly important to identify the three-dimensional location of antibody binding epitopes and to identify areas of the protein that cannot be changed in a vaccine due to structural constraints (Babon *et al.*, 2007, Bentley, 2006).

The structure of MSP1₁₉ from a number of species has been solved by Nuclear Magnetic resonance (NMR) or X-ray crystallography. There are advantages and disadvantages of using these two structural methods. NMR involves solving the structure of a protein in solution. This means that from NMR a group of structures is obtained because the protein is moving around in solution. In order to carry out NMR analysis large quantities of soluble purified protein must be produced. A concentration of around 1 mM is required to carry out NMR and this means the protein must also be stable at this concentration for long periods. The advantage of using NMR to solve the structure of a protein in solution is that the methodology can be used to look at the dynamics of the protein and to investigate kinetics of the protein binding to ligands. The quality of protein structures solved by NMR varies greatly and depends upon the quality of the spectra that are obtained. The overall resolution of protein structures solved by NMR is lower than for X-ray crystallography. The orientation of some side chains including asparagine, glutamine and threonine and the arrangement of hydrogen atoms can be distinguished by NMR but not by X-ray crystallography (Creighton, 1997, Freifelder, 1999). The main limitation of NMR is the size restriction. Even using labelled protein the upper size limit is around 50 KDa as beyond this size spectra become too complex and signals cannot be distinguished. There are newer techniques that have allowed larger 900 KDa protein structures to be solved but that is unusual for

NMR (Fernandez & Wider, 2003). X-ray crystallography involves crystallising a protein and using X-ray diffraction patterns to solve the structure. X-ray crystallography produces higher resolution structures than NMR and does not have an upper size limit. The structure of very large proteins and protein complexes such as a ribosome has been solved by X-ray crystallography. The main limitation of X-ray crystallography is the ability to crystallise the protein as not all proteins crystallise or crystallise in a biologically significant way. The quality of the structure also depends strongly on the quality of the crystal obtained (Creighton, 1997, Freifelder, 1999). Both methods can be used to complement one another and to provide different information about the protein. For example, NMR could be used to look at the dynamics of antibody binding to an antigen and identify the antibody binding site and X-ray crystallography could be used to crystallise the antibody-antigen complex and identify the contact surface between the antibody and antigen (Pizarro *et al.*, 2003).

1.5.1 Structure of MSP1₁₉

The structures of MSP1₁₉ from *P. falciparum* (Morgan *et al.*, 1999, Pizarro *et al.*, 2003), *P. vivax* (Babon *et al.*, 2007), *P. cynomolgi* (Chitarra *et al.*, 1999) and *P. knowlesi* (Garman *et al.*, 2003) have been solved by NMR or X-ray crystallography as described in table 1.1.

Table 1.1: Three-dimensional structures of MSP1₁₉ that have been solved by NMR or X-ray crystallography. This table is based on table 1 (Bentley, 2006)

Species	PDB entry	Method	Resolution	Year	Reference
<i>P. cynomolgi</i>	1B9W	X-ray	1.8 Å	1999	(Chitarra <i>et al.</i> , 1999)
<i>P. falciparum</i>	1CEJ	NMR		1999	(Morgan <i>et al.</i> , 1999)
<i>P. knowlesi</i>	1N1L	X-ray	2.4 Å	2002	(Garman <i>et al.</i> , 2003)
<i>P. falciparum</i> - Fab complex	1OB1	X-ray	2.9 Å	2003	(Pizarro <i>et al.</i> , 2003)
<i>P. vivax</i>	2NPR	NMR		2007	(Babon <i>et al.</i> , 2007)
The resolution is for X-ray crystal structures only.					

Figure 1.4 shows the backbone of MSP1₁₉ structures from the different species. The overall structure of MSP1₁₉ from the different *Plasmodium* species is very similar. The MSP1₁₉ proteins share common structural characteristics. The proteins are composed of two epidermal growth factor (EGF)-like motifs. An EGF-motif typically consists of 30 to 40 amino acids and has a characteristic disulphide bonding pattern that involves three intradomain disulphides (Wouters *et al.*, 2005). The main difference between the MSP1₁₉ structure of *P. falciparum* MSP1₁₉ and the other species is the disulphide bonding pattern of the first EGF domain. The cysteines in the two EGF-like domains of *P. falciparum* form the classic EGF disulphide bonding pattern of cysteine 1 to 3, 2 to 4 and 5 to 6. The cysteines in the second EGF domain of MSP1₁₉ from the other species form the same disulphide bonding pattern but the bonding pattern for the first EGF domain is different. In the first EGF-like domain, disulphide bond 2 to 4 is replaced by the side chain of tryptophan and a small residue (valine, isoleucine or threonine) (Babon *et al.*, 2007, Garman *et al.*, 2003) which pack together in the same volume as is observed for the disulphide bonds. The surface of MSP1₁₉ is charged and the charge distribution is unique for each protein (Babon *et al.*, 2007, Garman *et al.*, 2003).

1.6 Immunity to MSP1₁₉

1.6.1 *Evidence from animal studies of MSP1₁₉ as a vaccine candidate*

Animal studies have provided evidence that MSP1₁₉ has the potential to be a good malaria vaccine candidate. Daly *et al.* expressed MSP1₁₉ as a GST-fusion protein in *E. coli* and used it in immunisation studies in mice using Ribi adjuvant. The immunisation studies showed that both inbred and outbred mice that were immunised with the MSP1₁₉ GST construct were either partially or completely protected against challenge infection with lethal *P. yoelii* YM parasites. Daly *et al.* showed that the antibodies that were produced to MSP1₁₉ were able to react with native MSP1₁₉. In this study the GST portion of the fusion protein was required for the immune response to MSP1₁₉ because immunisation with MSP1₁₉ with the tag removed did not give protection. It was thought that the GST portion may have been playing a role in providing T-cell epitopes (Daly & Long, 1993). Ling *et al.* immunised BALB/c mice with MSP1₁₉ expressed in *E. coli* as a GST-fusion protein and without GST using Freund's complete adjuvant. The immunisation studies showed that MSP1₁₉ alone and GST-MSP1₁₉ protected the mice against challenge infection with *P. yoelii* YM. Analysis of the antibody titres of the mice showed that the highest antibody titres were seen in the mice that were protected. In this study, Ling *et al.* also demonstrated the importance of the disulphide bonds in the EGF domain structure for protection. Immunisation with reduced and alkylated MSP1₁₉ protein abolished the protective ability of the protein (Ling *et al.*, 1994). Later immunisation studies by Ling *et al.* using MSP1₁₉ preparations in adjuvants that have been developed for clinical use in humans gave protection that was as good as or better than the protection observed with Freund's complete adjuvant (Ling *et al.*, 1997). This suggests that the mouse malaria model can be used to test MSP1₁₉ with adjuvants relevant to humans. MSP1₁₉ from *P. yoelii* has also been shown to be protective in mice using oral vaccination (Zhang *et al.*, 2005).

Wan Omar *et al.* carried out immunisation studies with recombinant MSP1₁₉ from *P. berghei*. This study showed that when recombinant MSP1₁₉ in a formulation with alum

was used to immunise 10 mice 8 of the mice were protected from challenge infection with *P. berghei* (Wan Omar *et al.*, 2007).

de Koning-Ward *et al.* developed a rodent model to test immunity to *P. falciparum* (human malaria) MSP1₁₉ *in vivo*. In the model they used allelic replacement to create a *P. berghei* parasite that expressed the *P. falciparum* form of MSP1₁₉. In these studies mice were repeatedly exposed to the chimeric *P. berghei*-*P. falciparum* parasites and then challenged with homologous parasites. Analysis of the antibodies made to MSP1₁₉ by these mice indicated that high levels of MSP1₁₉-specific invasion inhibitory antibodies were produced. These studies showed that the level of these MSP1₁₉-specific invasion inhibitory antibodies correlated with the level of protection observed during challenge infection with blood stage chimeric parasites rather than the total titre of MSP1₁₉ specific antibodies (de Koning-Ward *et al.*, 2003). These studies therefore demonstrated *in vivo* the importance of the fine specificity of the antibody response to MSP1₁₉ for protection and the importance of understanding this response to engineering a successful vaccine.

1.6.2 Immunity to MSP1₁₉ in humans

There are conflicting studies about the association of antibody levels to *P. falciparum* MSP1₁₉ and their role in protection against malaria. There are a number of studies which show an association between antibodies to *P. falciparum* MSP1₁₉ and protection against malaria (Al-Yaman *et al.*, 1996, Branch *et al.*, 1998, Egan *et al.*, 1996, Hogg *et al.*, 1995, Riley *et al.*, 1992, Riley *et al.*, 1993). For example, Egan *et al.* carried out a study looking at antibody levels to *P. falciparum* MSP1₁₉ in children in Sierra Leone and Gambia. In these studies they found that there was a significant association between antibody response to MSP1₁₉ and resistance to clinical malaria infection. They calculated that antibodies to *P. falciparum* MSP1₁₉ in children in Sierra Leone provided approximately 40 % protection against clinical malaria (Egan *et al.*, 1996). Riley *et al.* carried out studies of naturally acquired immune responses to *P. falciparum* MSP1. These studies showed that high levels of antibodies were made to the MSP1₁₉ part and that these antibodies were associated with protection. The studies showed that the presence and levels of antibodies to MSP1 increased with age (Riley *et al.*, 1992, Riley

et al., 1993). These studies suggest that producing a vaccine that generates an appropriate antibody response to MSP1₁₉ has the potential to protect against malaria.

The studies of Dodoo *et al.* provide an argument that levels of antibodies to MSP1₁₉ do not correlate with protection. In their studies, Dodoo *et al.* looked at antibodies to MSP1₁₉ in Ghanaian children over an 18 month period. In these studies there was no difference in levels of antibodies with age and there was no evidence found of an association between the presence or level of antibodies against *P. falciparum* MSP1₁₉ and protection from malaria (Dodoo *et al.*, 1999). Phase I trials at the Walter Reed Army Institute of Research with an MSP1₁₉ construct called Falciparum Malaria Protein 1 in AS02A adjuvant showed that it was safe and immunogenic but when it was tested against challenge infection no protection was observed (Chiang *et al.*, 2006).

All the studies together suggest that protection from MSP1₁₉ is more complicated than the presence and level of the antibody response but the fine specificity of that response is more important in determining protection.

Studies by Nwuba *et al.* indicate the presence of three types of antibodies to MSP1₁₉ following natural immune responses to *P. falciparum* malaria. The antibodies are called: inhibitory, blocking and neutral (Nwuba *et al.*, 2002). Inhibitory antibodies inhibit erythrocyte invasion. In *in vitro* studies the inhibitory antibodies have been shown to inhibit secondary processing of MSP1 (Blackman *et al.*, 1994) and this is thought to be the mechanism by which they inhibit erythrocyte invasion. Blocking antibodies do not inhibit erythrocyte invasion or secondary processing. Blocking antibodies interfere with the activity of inhibitory antibodies by competing for binding to MSP1. Production of blocking antibodies can be induced by areas of MSP1₁₉ that are not the target of inhibitory antibodies and they can abolish the activity of inhibitory antibodies. This means that in the presence of inhibitory and blocking antibodies secondary processing and erythrocyte invasion can occur (Guevara Patino *et al.*, 1997). Blocking antibodies are thought to have arisen as an immune evasion mechanism by the parasite to avoid the activity of inhibitory antibodies. Neutral antibodies do not inhibit erythrocyte invasion or block inhibitory antibodies (Nwuba *et al.*, 2002). The presence of these three types of antibodies may explain why there is conflicting data on total antibody levels to MSP1₁₉ and protection. The studies suggest that understanding the fine specificity of the

inhibitory, blocking and neutral antibodies could help to develop a vaccine to specifically stimulate the production of inhibitory antibodies and therefore be more successful in producing the desired immune response than the native protein.

1.6.3 Studies on *P. falciparum* MSP1₁₉ antibody binding

Studies have been carried out on *P. falciparum* MSP1₁₉ to map the antibody binding sites. Dekker *et al.* formed complexes of MSP1₁₉ with 12.10 antibody (inhibitory monoclonal antibody) and 2F10 (neutral monoclonal antibody) and analysed them by electromicroscopy, analytical ultra centrifugation and dynamic light scattering. This analysis showed that the two antibodies formed a ring structure with MSP1₁₉. This study therefore provided evidence that the epitopes of the inhibitory antibody (12.10) and neutral antibody (2F10) had non-overlapping epitopes on MSP1₁₉ because in order to form the ring structure two antibodies with non-overlapping epitopes would be required. This study also suggested that the binding epitopes may be located on opposite sides of the molecule. Dekker *et al.* also carried out site directed mutagenesis studies making single amino acid changes to MSP1₁₉ and looking at the effect on binding to 12.10 and 2F10. The site directed mutagenesis studies showed that none of the amino acid changes affected the binding to both 12.8 and 2F10. The studies therefore suggested that the epitopes for binding to inhibitory and neutral antibodies were distinct (Dekker *et al.*, 2004).

Morgan *et al.* used Transverse Relaxation Optimised Spectroscopy (TROSY) NMR epitope mapping techniques to map the binding sites for one neutral (2F10) and two inhibitory antibodies (12.8 and 12.10). These studies indicated that there was a close relationship between the surface location of binding sites of the inhibitory antibodies that was distinct from the neutral antibody binding site (Morgan *et al.*, 2004). The studies of Morgan *et al.* and Dekker *et al.* have shown that the precise binding epitope for the antibodies is very important for their function. The position of the antibody binding epitope for 2F10 on the opposite side of the protein from the binding epitope of 12.8 and 12.10 may explain why the neutral antibody 2F10 does not interfere with the activity of the inhibitory antibodies (Dekker *et al.*, 2004, Morgan *et al.*, 2004). Morgan *et al.* also carried out cross saturation TROSY NMR studies to more precisely map the

antibody binding sites of 12.8 and 12.10. The studies map the antibody binding site to the first β -sheet of EGF domain 1 (Morgan *et al.*, 2005).

Uthaipibull *et al.* used site directed mutagenesis to alter individual amino acids in *P. falciparum* MSP1₁₉ and tested binding to inhibitory, blocking and neutral antibodies by western blotting and surface plasmon resonance. These studies highlighted individual amino acids that were important in binding to different types of antibodies. The study suggested that it may be possible to make MSP1₁₉ proteins for vaccination that bind to inhibitory antibodies but not to blocking antibodies by using a site directed mutagenesis approach to alter combinations of residues involved in blocking antibody binding. Uthaipibull *et al.* also used a PEPSCAN method to look at reactivity of octapeptides with the antibodies. PEPSCAN analysis of reactivity with 12.8 (inhibitory monoclonal antibody) and 1E1 (blocking monoclonal antibody) showed the antibodies have adjacent antibody binding sites (Uthaipibull *et al.*, 2001).

Understanding the three dimensional location of the antibody binding sites of inhibitory, neutral and blocking antibodies and the fine specificity of these antibodies will help in engineering effective proteins for vaccination.

1.6.4 Studies on *P. yoelii* MSP1₁₉ antibodies and antibody binding

Some of the residues found to be important for antibody binding are conserved between *P. falciparum* and the rodent parasite *P. yoelii* MSP1₁₉. *P. falciparum* and *P. yoelii* MSP1₁₉ have a high degree of sequence similarity with approximately 50 % sequence identity. Figure 1.5 shows an alignment of *P. yoelii* and *P. falciparum* MSP1₁₉ highlighting the EGF-like motif and conserved residues. Studies using *P. yoelii* MSP1₁₉ can be compared to those for *P. falciparum* to gain information about conservation across the species and identify areas of functional conservation.

Spencer *et al.* produced monoclonal antibodies against *P. yoelii* MSP1. They looked at the ability of the monoclonal antibodies to protect mice against blood stage challenge infection with *P. yoelii* YM by passive immunisation. The purified antibodies were injected into groups of mice at the time of parasite challenge and the development of

parasitaemia was monitored. They identified two monoclonal antibodies (B6 and F5) that mediate a substantial reduction in parasitaemia and all the mice in the groups injected with them cleared the parasite. They also identified two monoclonal antibodies (B10 and G3) that produced partial suppression of parasite growth. B6 and F5 are of the IgG3 subclass, B10 is of the IgG2b subclass and G3 is of the IgG1 subclass. B10 and G3 require both EGF domains of MSP1₁₉ to form their binding epitopes and F5 and B6 only require the first EGF domain to form their binding epitopes. Western blotting analysis of the antibody binding to reduced and alkylated GST-MSP1₁₉ have shown that the antibody binding epitopes are constrained by the disulphide bonds in MSP1₁₉. Spencer *et al.* carried out competition ELISA analysis to compare the binding epitopes of the antibodies. The ELISA studies showed that the binding epitopes for B10 and G3 may be identical as the epitopes overlap. The epitopes for B6 and F5 also overlap each other but are clearly distinct binding epitopes. The epitope for B6 and F5 are different from those of B10 and G3. The study therefore suggests there are a number of distinct antibody binding epitopes on MSP1₁₉ antibodies with the ability to suppress parasitaemia *in vivo* (Spencer Valero *et al.*, 1998).

Benjamin *et al.* identified differences in the sequence of MSP1₁₉ between *P. yoelii* isolates. They expressed MSP1₁₉ recombinant proteins in *E. coli* from the *P. yoelii* isolates. Binding studies of the recombinant proteins to monoclonal antibodies B6, F5, B10 and G3 were used to identify residues that could be important for antibody binding. The antibody binding studies showed that all of the monoclonal antibodies were able to bind to the MSP1₁₉ proteins from *P. yoelii* YM but none of the monoclonal antibodies were able to bind to the MSP1₁₉ proteins from *P. yoelii yoelii* 2CL and *P. yoelii nigeriensis* N67. Benjamin *et al.* were able to identify positions of residues that were changed in the isolates that bound to different monoclonal antibodies. Isolates that bound to none of the antibodies had changes to residues 12, 31, 41, 47 and 78 of MSP1₁₉. Benjamin *et al.* identified changes to position 16 (lysine to glutamic acid) and 17 (asparagine to histidine) that may affect binding of B6 and F5 monoclonal antibodies to MSP1₁₉. They also identified changes to residues 52 and 54 that may affect binding of B10 monoclonal antibodies to MSP1₁₉ (Benjamin *et al.*, 1999).

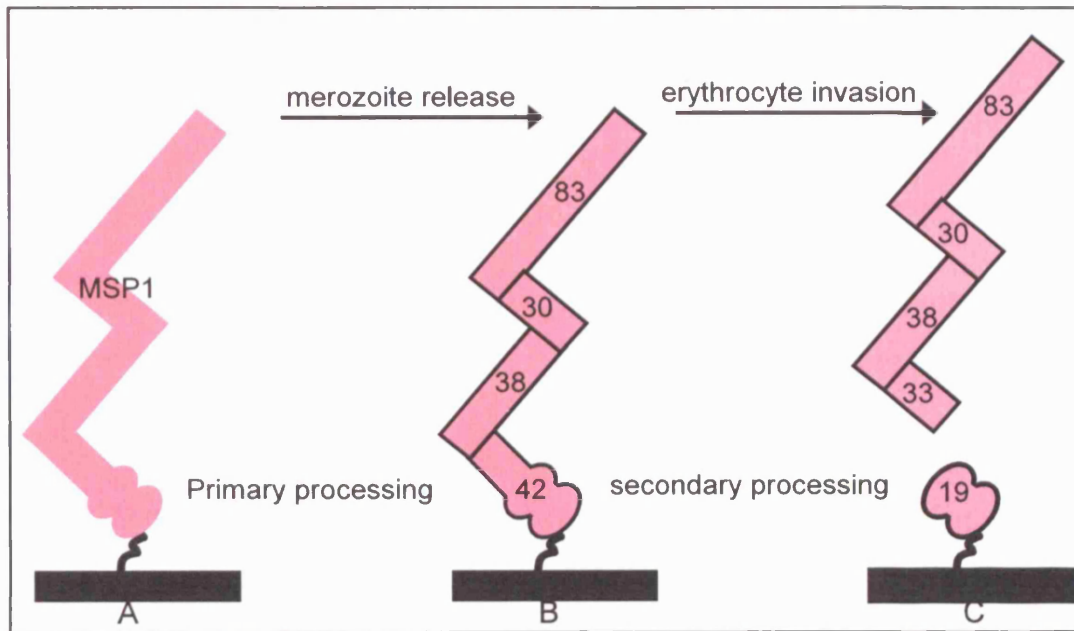


Figure 1.3: Schematic showing primary and secondary processing of MSP1.

A: Full length MSP1 before processing.

B: MSP1 following primary processing to form a complex of four polypeptides: MSP1₈₃ (83), MSP1₃₀ (30), MSP1₃₈ (38) and MSP1₄₂ (42).

C: MSP1 following secondary processing where MSP1₄₂ is cleaved into MSP1₃₃ (33) and MSP1₁₉ (19). MSP1₁₉ remains bound on surface of the merozoite and the rest of the complex is shed.

This figure is based on figure 1, Blackman *et al.* (Blackman *et al.*, 1994)

Figure 1.4: Comparison of the backbone of MSP1₁₉ from different *plasmodium* species.

The backbone of the best energy NMR structure of *P. falciparum* (Morgan *et al.*, 1999) MSP1₁₉ is shown with the first EGF domain in red and the second EGF domain in pink. The backbone of the best energy NMR structure of *P. vivax* (Babon *et al.*, 2007) MSP1₁₉ is shown with the first EGF domain in purple and the second EGF domain is green. The backbone of the crystal structure of *P. cynomolgi* (Chitarra *et al.*, 1999) MSP1₁₉ is shown with the first EGF domain in dark blue and the second EGF domain in light blue. The backbone of the crystal structure of *P. knowlesi* (Garman *et al.*, 2003) MSP1₁₉ is shown with the first EGF domain in orange and the second EGF domain in yellow. The N-terminus is shown in black and the C-terminus is shown in grey. The very ends of the protein are not defined in all the structures therefore the N and C-termini are the last residues shown in the structures.

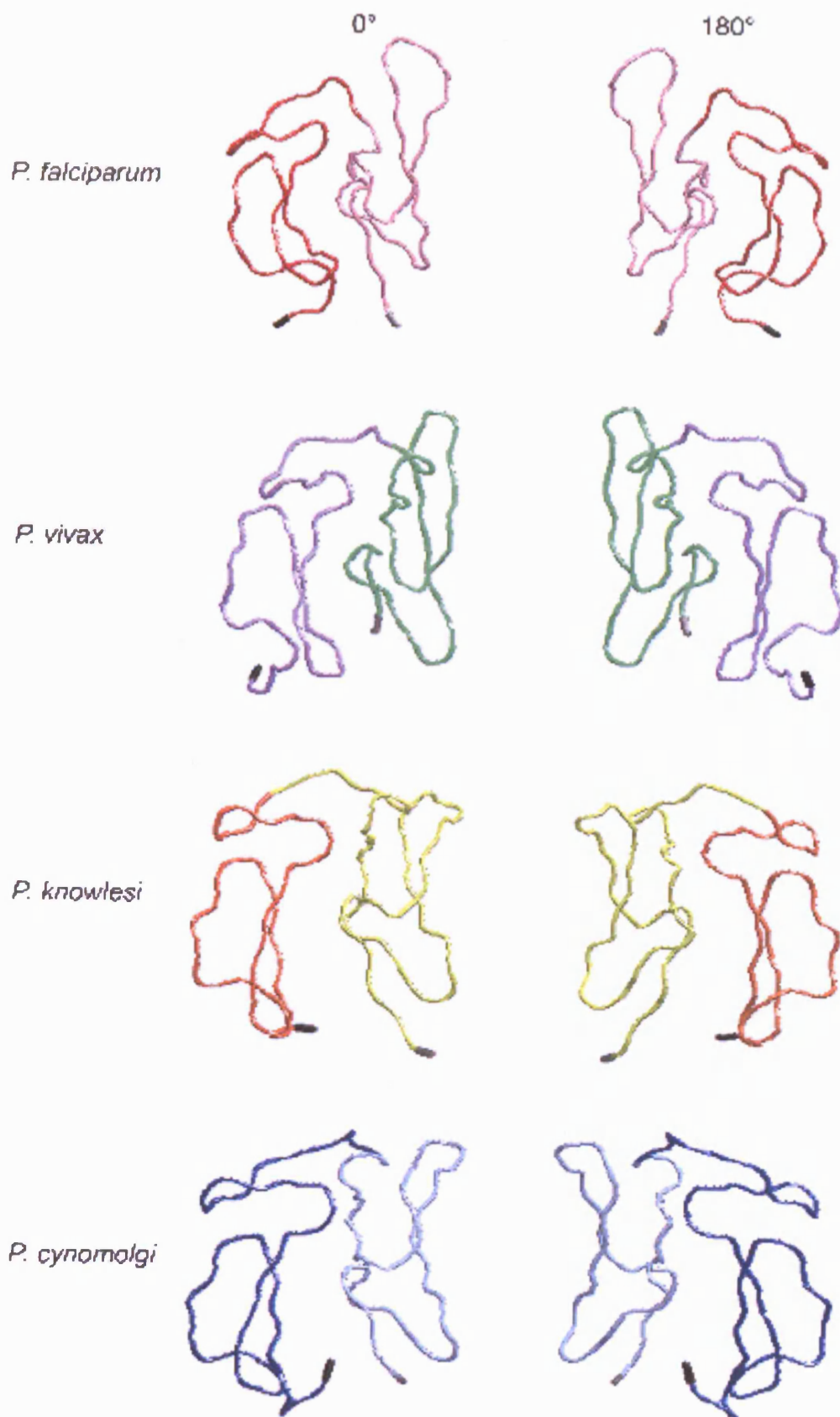




Figure 1.5: Alignment of *P. yoelii* (Py) and *P. falciparum* (Pf) MSP1₁₉ highlighting the conserved residues.

Conserved residues are highlighted in pink. Residues that are part of the conserved EGF motif are highlighted green. The alignment is based on figure 3 of Benjamin *et al.* (Benjamin *et al.*, 1999).

1.7 Aims of the project

The aim of this project is to map the antibody binding sites of *P. yoelii* MSP1₁₉. This will allow important residues for antibody binding to be determined. The data will be compared to the information in the literature for *P. falciparum* MSP1₁₉ inhibitory antibodies to identify common areas of antibody binding. This will indicate if there is a common mechanism of action for inhibitory antibodies across the species. This information could be important in developing antigens for vaccination to specifically stimulate the production of inhibitory antibodies. It could also be useful in directing research to understand the mechanism of action of inhibitory antibodies and their use as therapeutic agents.

The aim of the project will be broken into three main objectives:

- 1) A site directed mutagenesis approach will be used to identify important residues for antibody binding *in vitro* and this will be compared to *P. falciparum* data for inhibitory antibody binding.
- 2) The *P. yoelii* MSP1₁₉ variant proteins produced by site directed mutagenesis will be used in immunisation studies to determine whether any differences in *in vitro* antibody binding can correlate with differences in the ability of MSP1₁₉ to protect against challenge infection *in vivo*.
- 3) The structure of *P. yoelii* MSP1₁₉ and the MSP1₁₉ variant proteins will be examined by NMR to identify any differences in the structure caused by the amino acid changes and to identify if these residues also play a vital structural role. If these residues play a vital structural role they may be conserved across the species because of this role and may be unable to change due to immune pressure. This will help in identifying any residues that cannot be altered in vaccine development due to structural roles.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Buffers and solutions

PBS: 137 mM NaCl, 3 mM KCl, 8 mM Na₂HPO₄, 1.5 mM KH₂PO₄ in dH₂O

TAE (agarose gel running buffer): Prepared from a fifty times concentrated stock containing: 242 g TRIS base, 37.2 g Na₂EDTA.2H₂O and 57.1 ml glacial acetic acid per litre of water.

Oligonucleotides: synthetic oligonucleotides were all supplied by Sigma Genosys.

Restriction enzymes: All restriction enzymes were obtained from New England Biolabs or Roche.

Other general reagents: Chemical and general reagents used in this project were purchased from Sigma unless otherwise stated.

2.1.2 Bacterial cell culture: Bacteria and Plasmids

The following *E. coli* strains were used:

Strain	Supplier	Genotype
BL21(λ DE3)	Stratagene	F^- , <i>ompT</i> , <i>hsdS_B</i> , (<i>r_B</i> , <i>m_B</i> ⁻), <i>dcm</i> , <i>gal</i> , (λ DE3)
XL1-Blue Supercompetent cells	Stratagene	<i>recA1 endA1 gyrA96 thi-1</i> <i>hsdR17 supE44 relA1 lac</i> F' [<i>proAB</i> ⁺ <i>lac</i> F' <i>lac</i> Δ M15 <i>Tn10(tet')</i>]
One Shot TOP10 competent cells	Invitrogen	F^- <i>mcrA</i> Δ (<i>mrr-hsdRMS-</i> <i>mcrBC</i>) ϕ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74 recA1 araD139</i> Δ (<i>ara-leu</i>)7697 <i>galU galK</i> <i>rpsL</i> (Str ^R) <i>endA1 nupG</i>

LB: 1 % w/v bacto-tryptone, 0.5 % w/v bacto-yeast extract, 170 mM NaCl in dH₂O

LB-Agar: 1 % w/v bacto-tryptone, 0.5 % w/v bacto-yeast extract, 170 mM NaCl in dH₂O with 1.5 % w/v agar

Terrific broth: 12 g tryptone, 24 g yeast extract, 4 ml glycerol, 12.54 g K₂HPO₄, 2.31 g KH₂PO₄ per litre of water

SOC: 20 g Bacto tryptone, 5 g yeast extract, 0.584 g NaCl, 0.186 g KCl, 2.033 g MgCl₂.6H₂O, 2.464 g MgSO₄.7H₂O, 3.603 g glucose per litre of water

2.2 Methods: Binding studies

2.2.1 *Production of MSP1₁₉ variant clones as GST fusion proteins*

The DNA coding for the wildtype MSP1₁₉ of *P. yoelii* YM has been cloned into the *E. coli* expression vector pGEX3X to produce a GST fusion protein (GST-MSP1₁₉). The DNA for pGEX3X with MSP1₁₉ was prepared using the QIAprep Spin Miniprep kit (QIAGEN). QIAGEN miniprep kits are based on a modified alkaline lysis method (Birnboim & Doly, 1979) followed by the adsorption of the DNA to a silica based resin in the presence of high salt. Bacteria containing the desired plasmid are lysed under alkaline conditions and the lysate is subsequently neutralised and adjusted to high salt binding conditions. The use of a silica gel membrane optimises the adsorption of the DNA in mini-spin columns with the elution of RNA, proteins and metabolites. The DNA is subsequently eluted in water after a wash to remove endonucleases and salts.

Site-directed mutagenesis of wildtype GST-MSP1₁₉ was carried out using the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions to produce five variant proteins with single amino acid changes. Primers used in the mutagenesis are shown in table 2.1 with the bases involved in the amino acid changes highlighted in pink.

Table 2.1: Primers used in the Quikchange mutagenesis protocol

Variant	Primer 1	Primer 2
Arg12→Leu	5' GTGTTGATACACTAGATATTCCT AAAAATGCTGG3'	5' CCAGCATTTTTAGGAATATCTACT GTATCAACAC3'
Lys16→Glu	5' CAAGAGATATTCCTGAAAATGCT GGATGTTTTAG3'	5' CTAAAACATCCAGCATTTTAGGA ATATCTCTTG3'
Asn17→His	5' CAAGAGATATTCCTAAACATGCT GGATGTTTTAG3'	5' CTAAAACATCCAGCATCTTTAGGA ATATCTCTTG3'
Glu28→Lys	5' GAGATGATAATGGTACTAAGAA TGGAGATG3'	5' CATCTCCATTCTTAGTACCATTA TCATCTC3'
Glu28→Gln	5' GAGATGATAATGGTACTAAGAA TGGAGATG3'	5' CATCTCCATTCTTAGTACCATTA TCATCTC3'

Pilot mutagenesis studies were carried out to determine the optimum concentration of DNA template. This optimum concentration of 50 ng was used for all the mutagenesis experiments.

Following mutagenesis the DNA from a number of colonies was prepared using the QIAprep Spin Miniprep kit (QIAGEN) and sequenced at the Advanced Biotechnology Centre, Imperial College London or Cogenics to confirm the presence of the desired amino acid change. The following primers were used:

pGex seq F 5' CCAGCAAGTATATAGCATGG 3'

pGex seq R 5' CCGGGAGCTGCATGTGTACAG 3'

A double variant protein (Lys16→Glu and Glu28→Lys) was produced using Lys16→Glu variant DNA as a template and Glu28→Lys primers.

2.2.2 Expression of GST-MSP1₁₉ variants

Lysis buffer: 50 mM Tris/Cl pH 8.0, 1 mM EDTA, 0.2 % NP40

The GST-MSP1₁₉ variants and the GST-tag alone were all expressed using the same method. All expression steps were carried out at 37 °C. A single colony was used to inoculate 10 ml terrific broth-ampicillin (100 µg/ml) and grown for 8 hours. This culture was used to seed 100 ml terrific broth-ampicillin (100 µg/ml) and grown overnight. This culture was used to seed 500 ml terrific broth-ampicillin (100 µg/ml) and grown for 1 hour. The culture was then induced with 1 mM IPTG for 3 hours. Following induction, the bacterial cells were harvested by centrifugation at 3000 rpm, 4 °C for 10 minutes in Beckman J-6B centrifuge. The soluble fusion protein was recovered by resuspending the cells in 30 ml lysis buffer containing 1 mg/ml lysozyme (Sigma) and a complete protease inhibitor tablet (Roche). The cell pellets were incubated with shaking at 4 °C for 2 hours, with the addition of 10 µl of DNase I (Sigma) after 1 hour. The insoluble cell debris was removed by ultracentrifugation at 30,000 rpm (80,000 ×g), 4 °C for 45 minutes (Beckman L7 ultracentrifuge with 70Ti rotor). The supernatant was collected.

2.2.3 Purification of GST-MSP1₁₉ variants

Equilibration buffer: 50 mM Tris/Cl, 1 mM EDTA, 0.2 % v/v NP40 pH 8.0

Primary wash buffer: 50 mM Tris/Cl, 1 mM EDTA, 0.2 % v/v NP40 pH 8.0

Secondary wash buffer: 50 mM Tris/Cl, 1 mM EDTA pH 8.0

Elution buffer: 5 mM reduced glutathione, 50 mM Tris/Cl, 1 mM EDTA, pH 8.0

The protein was purified using glutathione agarose (Sigma) at 4 °C. The column was equilibrated with 5 column volumes of equilibration buffer. The supernatant was loaded onto the column. The column was washed first with 5 column volumes of primary wash buffer, followed by 5 column volumes of secondary wash buffer. The protein was eluted in 0.5 ml fractions with elution buffer and the fractions analysed by UV spectroscopy for the presence of protein. The protein fractions were dialysed against PBS using snakeskin pleated dialysis tubing MWCO 3,500 (Pierce). The protein samples were run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MOPS buffer (Invitrogen) and visualised by staining with Coomassie Brilliant Blue R-250 (Sigma). The protein was snap frozen and stored in small aliquots at -20 °C.

2.2.4 Quantification of GST-MSP1₁₉ variants

The concentration of the GST-MSP1₁₉ variants was estimated by measuring the absorbance of the proteins at 280 nm. The following formula was used to calculate the concentration: an absorbance reading of 1.0 at 280 nm (with a 1 cm path length) corresponds to a protein concentration of 0.5 mg/ml.

2.2.5 Western blotting

Samples of GST-MSP1₁₉ variants were diluted in 4x NuPAGE LDS sample buffer (Invitrogen). 10x NuPAGE reducing agent was added for western blotting with anti-GST antibody. The samples were run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MOPS buffer (Invitrogen) alongside dual colour protein standards (Biorad) and purified GST and BSA as controls. The proteins were transferred onto Protran BA nitrocellulose (Schleicher and Schuell Bioscience). The blot was blocked with 5 % Marvel milk in PBS. The nitrocellulose was incubated with various dilutions of antibodies diluted in PBS, 0.05 % Tween-20 (Sigma) as shown in table 2.2.

Table 2.2: Western blotting antibodies		
Antibody	Dilution	Incubation time
Anti-GST HRP conjugate (Amersham Biosciences)	1/5000	40 minutes
B6(Spencer Valero <i>et al.</i> , 1998)	2 µg/ml	2 hours
F5(Spencer Valero <i>et al.</i> , 1998)	10 µg/ml	2 hours
B10(Spencer Valero <i>et al.</i> , 1998)	2 µg/ml	2 hours
Goat anti-mouse IgG (H +L) HRP conjugate (Biorad) (Secondary antibody)	1/2000	20 minutes

The blots were washed following antibody incubation, 3 times with PBS, 0.05 % Tween-20 and once with PBS. The blot was visualised using the ECL kit (Amersham) according to the manufacturer's instructions.

2.2.5 *ELISA analysis of GST-MSP1₁₉ variant proteins*

All ELISA experiments were carried out with duplicate plates. All wash steps involved 3 washes with PBS, 0.05 % Tween-20 and 1 time with PBS. Initial pilot experiments were carried out to determine the optimum concentration of goat anti-GST antibody for coating the plate. In these experiments a range of concentrations of antibody from 0.2 to 10 µg/ml were used. The wells of 96-well flat-bottomed ELISA plates (Nunc Maxisorp) were coated with 100 µl of 0.4 µg/ml goat anti-GST antibody (Amersham) in sodium carbonate buffer (0.1 M, pH 9.6). The plates were incubated overnight at 4 °C and then washed. The wells were blocked with 50 µl per well of 1 % BSA in PBS. The plates were incubated at 37 °C for 1 hour and then washed. 100 µl of 1 µg/ml of GST-MSP1₁₉ variants in 1 % BSA in PBS were added to the desired wells. PBS and purified GST were used as controls. The plates were incubated for 2 hours at 37 °C and then washed. 100 µl of doubling dilutions of primary antibody (B6, F5 or B10 (Spencer Valero *et al.*, 1998)) in PBS were added to the wells. The plates were incubated for 1 hour at 37 °C and then washed. 100 µl of a 1/2000 dilution of secondary antibody (goat antimouse IgG (H+L) HRP, Biorad) in 1 % BSA in PBS was added to the wells. The plate was incubated for 30 minutes at 37 °C and then washed. In order to detect the peroxidase, 100 µl of freshly prepared o-phenylenediamine dihydrochloride (Sigma) in 0.05 M phosphate-citrate buffer was added to the wells. The plate was incubated at room temperature for 10 minutes before stopping the reaction with 50 µl of 1 M sulphuric acid. The absorbance was read at 490 nm.

2.2.6 *Surface plasmon resonance (SPR) analysis*

Measurements were performed on a BIAcore 2000 instrument (Biacore) at 25 °C. PBS, 0.05 % Tween-20 was used as a running buffer throughout. Anti-GST antibody was bound to the surface of a carboxymethyl dextran sensor chip CM5 using the GST capture kit and amine coupling kit (Biacore) according to the manufacturer's instructions. The GST-MSP1₁₉ variants were bound to the antibody using 70 µl of



solutions at 10 µg/ml diluted in PBS, 0.05 % Tween-20 for 7 minutes. After injecting, the response was allowed to stabilise for 3 minutes. The binding level was recorded. The binding assays were performed with monoclonal antibodies B6 (1.29 mg/ml), F5 (1.185 mg/ml) and B10 (1.18 mg/ml) (Spencer Valero *et al.*, 1998) diluted in PBS, 0.05 % Tween-20, at a constant flow rate of 5 µl min⁻¹ for 2 minutes. The binding level at steady state was recorded. The chip was washed with 40 µl of regeneration solution (10 mM glycine-HCl pH 2.2) for 2 minutes to wash off the GST-MSP1₁₉-antibody complex. Each binding assay was repeated in triplicate. Purified GST was used as a control.

2.2.7 Molecular modelling of *P. yoelii* MSP1₁₉

The Swiss Model Protein Modelling Server (<http://swissmodel.expasy.org>) (Guex & Peitsch, 1997, Peitsch, 1995, Schwede *et al.*, 2003) was used to create a homology model of *P. yoelii* MSP1₁₉. The “first approach method” was used. This method does a BLAST search for template sequences with 3D-structures in the Protein Databank (PDB) and makes a model based on these structures. The PDB structures used for the model were crystal and NMR structures of *P. falciparum* (PDB entries: 1ob1F (Pizarro *et al.*, 2003), 1ob1C (Pizarro *et al.*, 2003) and 1cejA (Morgan *et al.*, 1999)); *P. cynomolgi* (PDB entry 1b9wA (Chitarra *et al.*, 1999)) and *P. knowlesi* MSP1₁₉ (PDB entry 1n1iC (Garman *et al.*, 2003)). The model was displayed using Deepview/Swiss Pdb viewer (<http://www.expasy.org/spdbv>) (Guex & Peitsch, 1997, Peitsch, 1995, Schwede *et al.*, 2003) and manipulated using the RasTop programme (Valadon, 2007).

2.2.8 *In silico* variation of residues in *P. yoelii* MSP1₁₉

The amino acid variations made experimentally were made *in silico* on the protein model using the Deepview/Swiss Pdb viewer (Guex & Peitsch, 1997, Peitsch, 1995, Schwede *et al.*, 2003). The “mutate amino acids tool” was used which allows any amino acid in the model to be altered and the different possible rotamers to be displayed.

2.3 Methods: Immunisation Studies

All the animal handling was carried out by Sola Ogun and Madhu Kadekoppala. Mice were immunised with the GST-MSP1₁₉ variants to investigate their ability to protect against parasite challenge.

Groups of six 8-week-old BALB/c mice (from the Specific Pathogen Free unit, NIMR, London, UK) were immunised intraperitoneally (i.p.) with 10 µg of protein in Freund's Complete adjuvant (FCA). The response was boosted by a further two injections of 40 µg of protein in Freund's Incomplete adjuvant (FIA) 21 and 42 days later. The proteins used in the immunisation studies are described in table 2.3.

Table 2.3: Proteins used in immunisation studies	
Group	Protein
1	Wildtype GST-MSP1 ₁₉ (positive control)
2	Arg12→Leu GST-MSP1 ₁₉ variant
3	Lys16→Glu GST-MSP1 ₁₉ variant
4	Asn17→His GST-MSP1 ₁₉ variant
5	Glu28→Lys GST-MSP1 ₁₉ variant
6	Glu28→Gln GST-MSP1 ₁₉ variant
7	Lys16→Glu and Glu28→Lys double GST-MSP1 ₁₉ variant
8	GST (negative control)

The protein was prepared by mixing equal quantities of the protein diluted in PBS with FCA for the prime and FIA for the boost.

Serum samples were taken 14 days after the final immunisation. The blood was left for 30 minutes at room temperature to clot before centrifuging at $4722 \times g$ in a tabletop centrifuge to recover the serum. The serum samples were stored at -20 °C. The serum samples were used for ELISA analysis as described in section 2.3.1. The mice were

challenged 15 days after the final immunisation by intravenous (i.v) injection of 5×10^3 *P. yoelii* YM parasitized erythrocytes into the lateral vein of the tail. The parasitaemia was followed daily on blood films stained with 20 % Giemsa reagent (BDH).

The parasitaemia was measured by counting the number of parasite-infected cells on the Giemsa stained blood films under a microscope. At low levels of infection, 10 fields of view containing 200 cells were counted. Low levels of infection were classified as 2.5 % parasitaemia or less. At high levels of infection (more than 2.5 % parasitaemia), 1 representative field of view containing 200 cells was counted. The percentage parasitaemia was calculated as follows:

$$(\text{number of parasite-infected cells} \div \text{total number of cells}) \times 100$$

The infection was followed for at least 21 days and until the mice had cleared the parasites. The mice were considered to have cleared the parasites if no parasites were detected on at least 3 consecutive days.

2.3.1 ELISA analysis of serum samples from immunisation studies

All ELISA experiments were carried out with duplicate plates. All wash steps involved 3 washes with PBS, 0.05 % Tween-20 and 1 time with PBS. The wells of 96-well flat-bottomed ELISA plates (Nunc Maxisorp) were coated with 100 μ l of 1 μ g/ml of MSP1₁₉ his-tagged variants (production and purification of his-tagged proteins described in section 2.4.17) in sodium carbonate buffer (0.1 M, pH 9.6). The plates were incubated overnight at 4 °C and then washed. The wells were blocked with 50 μ l per well of 1 % BSA in PBS. The plates were incubated at 37 °C for 1 hour and then washed. 100 μ l of doubling dilutions of serum samples in PBS were added to the wells. The plates were incubated for 1 hour at 37 °C and then washed. 100 μ l of a 1/2000 dilution of goat antimouse IgG (H+L) HRP antibody (Biorad), in 1 % BSA in PBS was added to the wells. The plate was incubated for 30 minutes at 37 °C and then washed. In order to detect the peroxidase, 100 μ l of freshly prepared o-phenylenediamine dihydrochloride (Sigma) in 0.05 M phosphate-citrate buffer was added to the wells. The plate was incubated at room temperature for 10 minutes before stopping the reaction with 50 μ l of 1 M sulphuric acid. The absorbance was read at 490 nm.

2.4 Methods: Protein preparation for structural NMR studies

2.4.1 *Production of MSP1₁₉ variant clones as his-tagged proteins*

The MSP1₁₉ variants were produced as his-tagged proteins for use in structural NMR studies (as described in section 2.5).

2.4.2 *Production of recodonised MSP1₁₉ variant genes*

A construct for the MSP1₁₉ variant genes for expression in *Pichia pastoris* as a hexahis-tagged protein was designed as shown in figure 2.1. The N-glycosylation site NGT was changed to DGT.

The MSP1₁₉ variant genes were recodonised for optimum expression in *Pichia pastoris* by GENEART. The genes were supplied cloned into the GENEART standard vector and lyophilised.

2.4.3 *Preparation of DNA for electroporation into Pichia pastoris*

The scheme used for the preparation of DNA for electroporation into *P. pastoris* is illustrated in figure 2.2. The methods used in this scheme are described below.

2.4.4 *Restriction enzyme digestion*

Double restriction digests with *Sna*BI (5 units/μl) and *Avr*II (4 units/μl) of pPIC9K vector (Invitrogen) and GENEART standard vector containing MSP1₁₉ variant genes were carried out to prepare them for ligation.

Typical restriction digest conditions used were:

3 ug DNA

1 μl Restriction enzyme 1

1 μ l Restriction enzyme 2

3 μ l 10 x BSA (10 mg/ml)

3 μ l Restriction enzyme buffer

Distilled water up to a final reaction volume of 30 μ l.

Incubation at 37 °C for 1 hour 30 minutes.

2.4.5 *Dephosphorylation of DNA*

Calf intestinal alkaline phosphatase (CIP) (Roche) was used to remove 5' terminal phosphate groups from vector DNA prior to ligation reactions, to reduce the likelihood of vector re-ligation. 2 units of CIP were added to the digested pPIC9K vector DNA for 1 hour 40 minutes at 37 °C followed by incubation for 15 minutes at 50 °C. The CIP was inactivated following dephosphorylation by incubation at 72 °C for 10 minutes.

2.4.6 *Agarose gel electrophoresis*

DNA loading buffer: 0.25 % bromophenol blue, 0.25 % xylene cyanol FF and 30 % glycerol in water.

DNA markers: 10 μ l of Quick-Load 1 kb DNA ladder (Biorad) and Quick-Load 100 bp DNA ladder (Biorad) per gel.

Agarose gel electrophoresis was used for isolation of DNA fragments, DNA agarose gel extraction, quantification of DNA and analysing DNA. Agarose gels were made by dissolving agarose in 1 \times TAE buffer to a final concentration of 1 % agarose w/v. The constituents were melted and poured after cooling and the addition of ethidium bromide (Biorad) to 0.4 μ g/ml into Anachem gel tanks. Gels were run at 40 mA for the required time and the DNA bands were visualised using the UV transilluminator. The concentration of DNA was estimated by comparisons of the apparent brightness of the ethidium bromide stained sample DNA with the DNA markers.

2.4.7 Agarose gel extraction

Digested vector and insert DNA and ligated DNA was purified by gel extraction. DNA was extracted from agarose gels by using the QIAquick gel extraction kit (QIAGEN) according to the manufacturers' instructions. The DNA was eluted with 30 µl sterile distilled water and subsequently quantified by agarose gel electrophoresis.

2.4.8 Ligation reactions

Digested recodonised MSP1₁₉ variant genes were ligated into digested and dephosphorylated pPIC9K vector DNA using T4 DNA ligase (Roche). The following conditions were used for ligation reactions:

50 ng Vector DNA

50 ng Insert DNA

1.1 µl T4 DNA ligase buffer

1 µl T4 DNA ligase (1 unit/µl)

Distilled water up to a final reaction volume of 11 µl

Reactions incubated for 18 hours at 16 °C. The ligase was inactivated following the ligation by incubation at 65 °C for 10 minutes.

2.4.9 Transformation of chemically competent *E. coli*

2 µl ligation reactions (described in section 2.4.8) were added to thawed 50 µl aliquots of One Shot TOP10 competent cells. After 30 minutes on ice the cells were heat shocked in a 42 °C water bath for exactly 30 seconds. After a further 2 minutes on ice, 250 µl of pre-warmed SOC medium was added to each aliquot and incubated shaking (225 rpm) at 37 °C for 1 hour. All of the aliquots were individually spread on prepared LB-agar plates containing 50 µg/ml ampicillin to get single colonies. The LB-agar plates were incubated at 37 °C overnight.

Following the transformation, DNA from a number of colonies was prepared using the QIAprep Spin Miniprep kit (QIAGEN) and sequenced at the Advanced Biotechnology

Centre, Imperial College or Cogenics to confirm the presence of the recodonised MSP1₁₉ genes and to confirm there had been no mutations during the DNA preparation.

The following primers were used:

3' AOX1 5' GCAAATGGCATTCTGACATCC 3'

a-factor 5' TACTATTGCCAGCATTGCTGC 3'

2.4.10 Production of E. coli glycerol stocks

The colonies were also used to inoculate 2 ml LB and grown shaking overnight at 37 °C. Following overnight growth, the cultures were divided into 500 µl aliquots and sterile glycerol was added to a final concentration of 15 % v/v. The cultures were snap frozen and stored at -80 °C.

2.4.11 Large scale DNA purification

The glycerol stocks of Top10 cells containing the pPIC9K with the recodonised MSP1₁₉ genes were used to inoculate cultures to produce DNA for purification using the QIAGEN HiSpeed Plasmid Midi Kit (QIAGEN) according to the manufacturers' instructions. The DNA was eluted in 500 µl sterile distilled water.

2.4.12 Production of Glu28→Gln his-MSP1₁₉ variant using Quikchange XL site-directed mutagenesis kit

Site-directed mutagenesis of wildtype his-MSP1₁₉ was carried out using the QuikChange XL site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions to produce Glu28→Gln his-MSP1₁₉ variant. The following primers were used:

E28Qo1: 5' GAGATGACGACGGTACTCAAGAGTGGAGATGTTTGTTC 3'

E28Qo2: 5' CAACAAACATCTCCACTCTTGAGTACCGTCGTCATCTC 3'

Following the mutagenesis, DNA from a number of colonies was prepared using the QIAprep Spin Miniprep kit (QIAGEN) and sequenced at Cogenics to confirm the presence of the desired amino acid change. The following primers were used:

3'AOX1 5' GCAAATGGCATTCTGACATCC 3'
a-factor 5' TACTATTGCCAGCATTGCTGC 3'

2.4.13 Linerisation of pPIC9K for transformation into *Pichia pastoris*

The purified pPIC9K with the recodonised MSP1₁₉ genes was linerised for transformation into *P. pastoris* by restriction digestion (as described in section 2.4.4) with *SacI* (20 units/μl).

2.4.14 Preparation of electrocompetent *Pichia pastoris*

The following *P. pastoris* strain was used: GS115 (*his4*, Mut⁺, Invitrogen)

YND: 0.67 % Yeast Nitrogen Base without amino acids (BD Difco), 1 % glucose

Electroporation buffer: 270 mM sucrose, 10 mM Tris-HCl pH 7.5, 1 mM MgCl₂

YEPD: 1 % yeast extract, 2 % peptone, 2 % glucose

MD agar: 1.34 % YNB (with ammonium sulphate without amino acids), 4×10^{-5} biotin, 2 % dextrose with 1.5 w/v agar

A GS115 stab (Invitrogen) was used to inoculate 5 ml YEPD medium and grown shaking (250 rpm) at 30 °C overnight. This culture was diluted 100 fold in 100 ml YEPD and grown shaking (250 rpm) at 30 °C overnight to an optical density at 600 nm of 1.2-1.5. The yeast cells were harvested by centrifugation at $3000 \times g$ 4 °C for 10 minutes in Beckman J-6B centrifuge. Following centrifugation, the cells were resuspended in 20 ml 50 mM potassium phosphate pH 6.0, 25 mM DTT. The cells were incubated for 15 minutes at 30 °C. Following the incubation, the cells were harvested by

centrifugation at $3000 \times g$ $4^\circ C$ for 10 minutes. The cells were washed with electroporation buffer as shown in the table below. All wash steps were carried out on ice. The cells were harvested by centrifugation at $3000 \times g$ $4^\circ C$ for 10 minutes in between washes and gently resuspended.

Wash 1	100 ml electroporation buffer
Wash 2	50 ml electroporation buffer

Following the washes, the cells were resuspended in 500 μl electroporation buffer to give approximately 2×10^{10} cells/ml. 60 μl aliquots of cells into transferred into pre-chilled 2 mm electroporation cuvettes (Biorad). The cells were mixed with 1 μg of recodonised MSP1₁₉ variant genes in pPIC9K. For the electroporation a Biorad Gene Pulser was used. The Gene Pulser was set at a 1.5 kV pulse with a 25 μF capacitor and the pulse controller set at 400 Ω . After the pulse the time constant was recorded (7.4 ms indicated a successful electroporation) and 1ml YEPD was immediately added to the cuvettes at room temperature. The cells were incubated at $30^\circ C$ for 1 hour. Following the incubation, the cells were transferred to 15 ml falcon tubes for washing to remove the YEPD medium. The cells were centrifuged at $2000 \times g$ for 10 minutes and then washed with 5 ml YND. The cells were harvested by centrifugation at $2000 \times g$ for 10 minutes and then resuspended in 1 ml YND for plating. 250 μl of cells were plated onto MD agar plates and incubated at $30^\circ C$ for three days. The remaining electroporation mixture was stored at $4^\circ C$ for plating if there was no growth on the MD agar plates.

2.4.15 Geneticin screening of transformants

YEPD-agar: 1 % yeast extract, 2 % peptone, 2 % glucose with 2 % w/v agar

3 ml sterile distilled water was added to the transformant plates. The colonies were gently resuspended with a sterile plastic spreader. The cell suspension was transferred to a sterile falcon tube. The OD at 600 nm was recorded to determine the concentration of cells. Serial dilutions were carried out until at OD at 600 nm of 1 was obtained. This was equal to 5×10^7 cells/ml. 40 μl cells at $OD_{600\text{ nm}} = 1$ were added to 960 ml sterile

distilled water to give a 10^5 cells per 100 μ l. 100 μ l cells were spread onto YEPD agar plates containing the following concentrations of geneticin (Geneticin selective antibiotic liquid, Invitrogen): no geneticin, 0.25 mg/ml, 0.5 mg/ml, 1 mg/ml and 2 mg/ml. The plates were incubated for up to 1 week at 30 °C until colonies appeared.

4 large colonies from each 2 mg/ml plate were re-streaked onto fresh YEPD agar plates containing 2 mg/ml geneticin to ensure that these colonies could definitely grow at 2 mg/ml geneticin concentrations. Each of the colonies was used to inoculate 2 ml cultures of YEPD and grown overnight shaking (250 rpm) at 30 °C. Following overnight growth, the cultures were divided into 500 μ l aliquots and sterile glycerol was added to a final concentration of 15 % v/v. The cultures were snap frozen and stored at -80 °C.

2.4.16 Small scale expression time course in *P. pastoris*

BMGlc: 100 mM potassium phosphate pH 6.0, YNB (0.34 % w/v) (BD Difco: YNB without amino acids and without ammonium sulphate) ammonium sulphate (0.2 % w/v), biotin (4×10^{-5} biotin, w/v), Sigma antifoam 289 (0.01 % v/v), D-glucose (0.5 % w/v)

BMMY: 100 mM potassium phosphate pH 6.0, YNB (0.34 % w/v) (BD Difco: YNB without amino acids and without ammonium sulphate) ammonium sulphate (0.2 % w/v), biotin (4×10^{-5} biotin, w/v), Sigma antifoam 289 (0.01 % v/v), methanol (1 % v/v)

All expression steps were carried out at 30 °C with shaking (250 rpm). A single colony of cells from the 2 mg/ml geneticin plates were used to inoculate 4 ml BMGlc and grown for 24 hours. This culture was used to seed 100 ml BMGlc and grown for approximately 67 hours. The yeast cells were harvested by centrifugation at $2000 \times g$ for 5 minutes at room temperature in Beckman J-6B centrifuge. The supernatant was removed and the cells were resuspended in 50 ml BMMY. The methanol in the BMMY was used to induce the expression of the protein using the AOX1 gene. The cells were transferred to 250 ml shake flasks and incubated for 96 hours. 100 % methanol was added to the cultures to a final concentration of 1 % every 24 hours to replenish the methanol that had been used up making the protein. 1 ml aliquots were removed from

the flasks every 24 hours. The aliquots were centrifuged at 13,000 rpm in a tabletop microcentrifuge for 2 minutes at room temperature. The supernatant was transfer to a separate tube for analysis. The supernatant and cell pellet were stored at -80 °C until required for analysis. After 96 hours the remaining cell culture was centrifuged at 2000 × g for 10 minutes at room temperature. The supernatant was removed and stored at -80 °C until required. For analysis, the supernatant aliquots from the time course were concentrated 10 times using the vivaspin 5K MWCO concentrator (Vivaspin). The supernatant samples were then run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer (Invitrogen) and the protein was visualised by staining with Coomassie Brilliant Blue R-250 (Sigma).

2.4.17 Large scale expression of his-tagged MSP1₁₉ variants

The labelled ¹⁵N labelled ammonium sulphate and ¹³C labelled glucose and ¹³C labelled methanol were supplied by Cambridge Isotope Laboratories Inc as follows: ¹³C methanol 99%, ¹⁵N ammonium sulphate 99% and D-glucose (U-13C6) 99%.

Large scale expression of the all of the his-MSP1₁₉ variants was carried out with ¹⁵N labelled ammonium sulphate to make ¹⁵N labelled protein for 1-Dimensional and 2-Dimensional Nuclear Magnetic Resonance (NMR) analysis. In addition, large scale expression of wildtype his-MSP1₁₉ and Glu28→Lys his-MSP1₁₉ variant were carried out with ¹⁵N labelled ammonium sulphate, ¹³C labelled glucose and ¹³C labelled methanol to make doubly labelled ¹⁵N and ¹³C labelled protein for 3-Dimensional NMR analysis. All expression steps were carried out at 30 °C with shaking (250 rpm). The glycerol stocks of cells for each variant were used to inoculate 4 × 5 ml BMGlc and grown for 24 hours. These cultures was used to seed 150 ml BMGlc and grown for approximately 67 hours. The yeast cells were harvested by centrifugation at 2000 × g for 10 minutes at room temperature in Beckman J-6B centrifuge. The supernatants were removed and the cell pellets were washed with 10 ml methanol. The cells pellets were resuspended in 4 × 500 ml BMMY in 2 L shake flasks and incubated for 72 hours to express the protein. 100 % methanol was added to the cultures to a final concentration of 1 % every 24 hours to replenish the methanol that had been used up making the protein. The cells were harvested by centrifugation at 2000 × g for 10 minutes at room

temperature in Beckman J-6B centrifuge. The supernatants were transferred to fresh tubes. The cell pellets were frozen at -80 °C for analysis if there was no protein in the supernatants. Complete EDTA-free protease inhibitor tablets (Roche) were added to the supernatants. The supernatants were concentrated using the Quick Stand benchtop system (Amersham Biosciences). The supernatants were first passed through a 0.45 micron filter to remove any cell debris followed by concentration down to 100 ml using a 5,000 Da MWCO hollow fibre filter.

2.4.18 Large scale purification of his-tagged MSP1₁₉ variants

Equilibration buffer: 50 mM NaPO₄, 300 mM NaCl, pH 7.2

Primary wash buffer: 50 mM NaPO₄, 300 mM NaCl, 10 mM imidazole, pH 7.2

Secondary wash buffer: 50 mM NaPO₄, 300 mM NaCl, 20 mM imidazole, pH 7.2

Elution buffer: 50 mM NaPO₄, 300 mM NaCl, 250 mM imidazole, pH 7.2

The proteins were purified using a batch purification method. All centrifugation steps were carried out at 2000 × g for 10 minutes at room temperature. Sodium chloride was added to the concentrated supernatants to a final concentration of 300 mM. The pH of the concentrated supernatant was increased to pH 7.5 using sodium hydroxide. The concentrated supernatant was centrifuged to remove any precipitate. 4 ml Ni-NTA agarose (QIAGEN) was washed twice with equilibration buffer and centrifuged between each wash. The concentrated supernatant was mixed with the Ni-NTA agarose and put onto a rotary shaker overnight at 4 °C to allow maximum binding. The agarose and concentrated supernatant was centrifuged to collect the Ni-NTA agarose. The supernatant was removed leaving a pellet of Ni-NTA agarose. The Ni-NTA agarose was resuspended in 5 ml of concentrated supernatant. The Ni-NTA agarose was loaded into a 10 ml disposable plastic column (Pierce) and allowed to settle. The column was washed with 4 column volumes of primary wash buffer. The column was then washed with 4 column volumes secondary wash buffer. The protein was eluted in 1 ml fractions with elution buffer and the fractions analysed by UV spectroscopy for the presence of

protein. The protein fractions were dialysed against double distilled water using snakeskin pleated dialysis tubing MWCO 3,500 (Pierce). The purified protein samples were then run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer (Invitrogen) and the protein was visualised by staining with Coomassie Brilliant Blue R-250 (Sigma).

2.4.19 Quantification of his-MSP1₁₉ variants

The concentration of the his-MSP1₁₉ variants was determined by measuring the absorbance of the proteins at 280 nm. The extinction coefficient and molecular weight of the wildtype his-MSP1₁₉ was determined using the ExPASy ProtParam tool (Gill & von Hippel, 1989).

Figure 2.1: his-MSP1₁₉ variants recodonised gene sequence

A: The recodonised gene sequence for wildtype his-MSP1₁₉ created by GENEART. The his-tag is in pink, the factor Xa cleavage site in blue, the residue 12, 16, 17, 28 variations are highlighted in red with the codons that are different in the variants are written in red beneath the residues. The residue highlighted in green has been changed from N in the wildtype to remove the N-glycosylation site.

B: Schematic representation of his-MSP1₁₉ that was cloned into the pPIC9K vector.

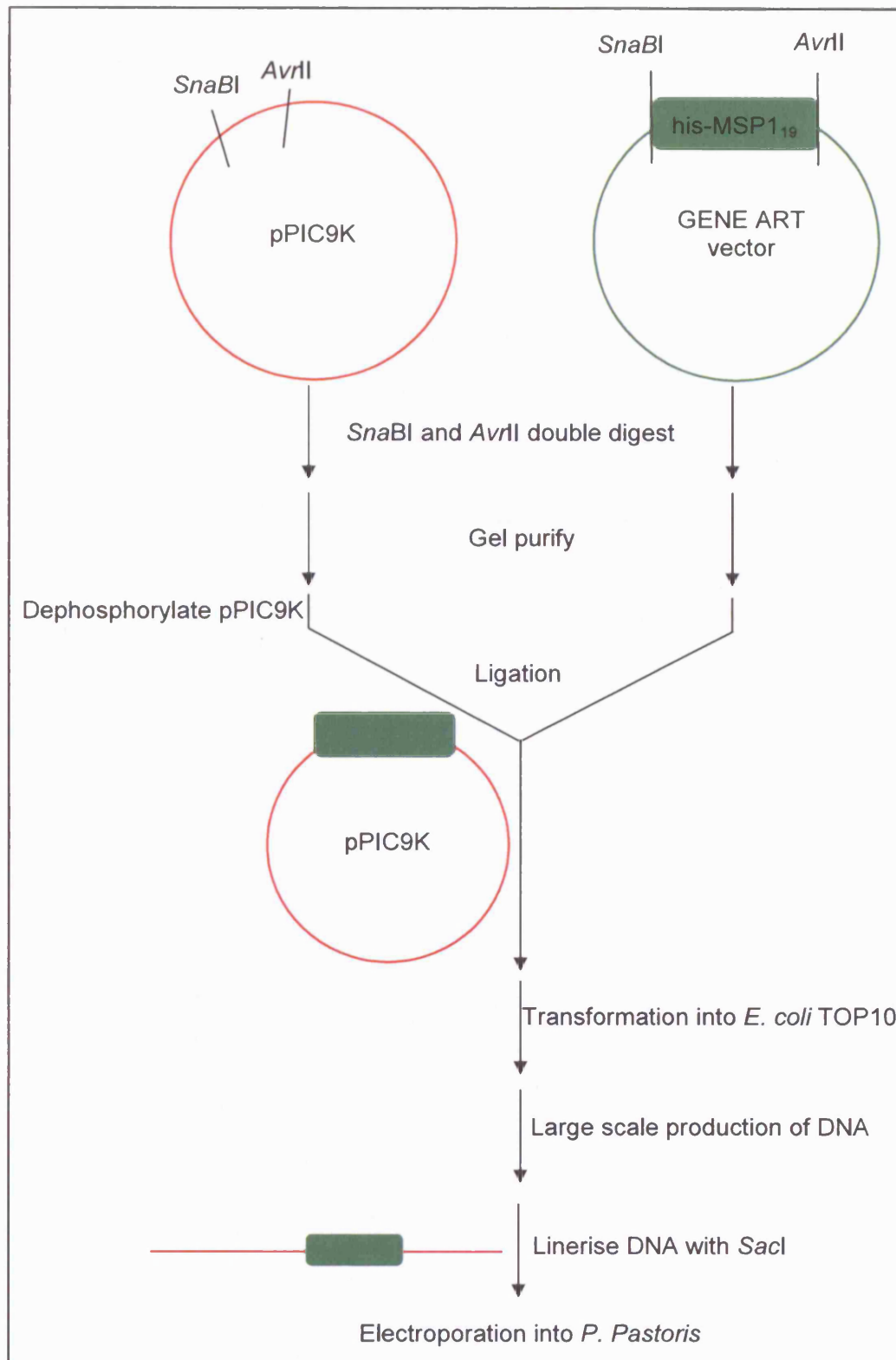
A

1
GGTACCTACGTACATCATCACCACCACCACATTGAAGGTAGAGGTGTTGACCCAAAGCAT
-----+-----+-----+-----+-----+-----+-----+
CCATGGATGCATGTAGTAGTGGTGGTGGTGTAACTTCCATCTCCACAACCTGGGTTTCGTA
 H H H H H H I E G R G V D P K H
61
GTTTGTGTTGACACTAGAGACATCCCAAAGAACGCTGGTTGTTTCAGAGATGACGACGGT
-----+-----+-----+-----+-----+-----+-----+
CAAACACAACCTGTGATCTCTGTAGGGTTTCTTGCGACCAACAAGTCTCTACTGCTGCCA
V C V D T R D I P K N A G C F R D D D G
 TTG GAA CAT
121
ACTGAAGAGTGGAGATGTTTGTGGGTTACAAGAAGGGTGAAGGTAACACTTGTGTTGAG
-----+-----+-----+-----+-----+-----+-----+
TGACTTCTCACCTCTACAAACAACCCAATGTTCTTCCCACTTCCATTGTGAACACAACCTC
T E E W R C L L G Y K K G E G N T C V E
 AAG
181
AACAACAACCCCAACTTGTGACATCAACAACGGTGGTTGTGATCCAACCTGCTTCCTGTCAA
-----+-----+-----+-----+-----+-----+-----+
TTGTTGTTGGGTTGAACACTGTAGTTGTTGCCACCAACACTAGGTTGACGAAGGACAGTT
N N N P T C D I N N G G C D P T A S C Q
241
AACGCTGAGTCTACTGAGAACTCCAAGAAGATCATCTGTACTTGTAAAGAGCCAACTCCA
-----+-----+-----+-----+-----+-----+-----+
TTGCGACTCAGATGACTCTTGAGGTTCTTCTAGTAGACATGAACATTTCTCGGTTGAGGT
N A E S T E N S K K I I C T C K E P T P
301
AACGCTTACTACGAGGGAGTTTTCTGTTCTTCTTCTTCTTAGTAGCCTAGGAGCTC
-----+-----+-----+-----+-----+-----+-----+
TTGCGAATGATGCTCCCTCAAAAGACAAGAAGAAGAATCATCGGATCCTCGAG
N A Y Y E G V F C S S S S * *

B



Figure 2.2: Overview of the strategy used to prepare the pPIC9K and his-MSP1₁₉ recodonised genes for transformation into yeast.



2.5 Methods: Nuclear Magnetic Resonance (NMR) Spectroscopy

The NMR spectroscopy experiments and processing of the spectra was carried out by Berry Birdsall. The analysis and assignment of the NMR spectra was carried out in collaboration with Berry Birdsall. The theory behind the NMR techniques and schematic examples of spectra are in chapters 6 and 9.

2.5.1 *Preparation of his-MSP1₁₉ variants for NMR spectroscopy*

NMR buffer: 50 mM potassium chloride, 25 mM potassium phosphate, pH 6.5.

5 mg of the his-MSP1₁₉ variants were snap frozen and freeze dried in Eppendorf tubes. The samples were stored at -20 °C until required. The samples were resuspended in 360 µl NMR buffer and 20 µl deuterium oxide was added to give a protein sample at 1 mM. The samples were centrifuged at 13,000 rpm in a tabletop microcentrifuge for 1 minute to remove any precipitate. The samples were transferred to shegemi tubes for NMR analysis.

2.5.2 *¹⁵N Heteronuclear Single Quantum Correlation (¹⁵N-HSQC) Spectroscopy*

All NMR spectra were recorded at 25 °C (unless otherwise stated) at 600, 700 or 800 MHz on Varian or Bruker NMR spectrometers. Water suppression was carried out using the WATERGATE sequence (Piotto *et al.*, 1992). ¹⁵N-HSQC spectra were acquired for all ¹⁵N labelled his-MSP1₁₉ variants. The spectra were processed using NMRpipe and NMRDraw (Delaglio *et al.*, 1995) and analysed using SPARKY software (Goddard & Kneller). The spectra for the his-MSP1₁₉ variants was overlayed on top of the wildtype his-MSP1₁₉ spectrum in SPARKY to compare the positions of the peaks in the spectra to identify areas that had changed.

2.5.3 Using NMR to determine the 3D structure of wildtype his-MSPI₁₉ and Glu28→Lys his-MSPI₁₉

All NMR spectra were recorded at 25 °C or 35 °C at 600, 700 or 800 MHz on Varian or Bruker NMR spectrometers. Water suppression was carried out using the WATERGATE sequence (Piotto *et al.*, 1992). NMR spectra were acquired for ¹⁵N labelled proteins and doubly labelled ¹³C/¹⁵N labelled proteins. The spectra were processed using NMRpipe and NMRDraw (Delaglio *et al.*, 1995) and analysed using SPARKY (Goddard & Kneller) and Xeasy (Bartels *et al.*, 1995) software. The NMR spectra were analysed to assign the backbone residues, side chain atoms and long distance restraints as shown in table 2.4.

Table 2.4: NMR spectroscopy experiments used in 3D structure determination

Spectrum name	Sample	Use
¹⁵ N -HSQC	¹⁵ N labelled protein	Assigning backbone atoms
¹³ C -HSQC	¹⁵ N/ ¹³ C labelled protein	Assigning backbone atoms and side chain atoms
HNCO	¹⁵ N/ ¹³ C labelled protein	Assigning backbone atoms
HNCACB	¹⁵ N/ ¹³ C labelled protein	Assigning backbone atoms
CBCACONH	¹⁵ N/ ¹³ C labelled protein	Assigning backbone atoms
HCCCONH	¹⁵ N/ ¹³ C labelled protein	Assigning side chain atoms
HCCH-TOCSY	¹⁵ N/ ¹³ C labelled protein	Assigning side chain atoms
¹⁵ N –NOESY-HSQC	¹⁵ N labelled protein	Identifying short and long distance restraints
¹³ C–NOESY-HSQC	¹⁵ N/ ¹³ C labelled protein	Identifying short and long distance restraints
¹³ C –HSQC tuned to aromatic region	¹⁵ N/ ¹³ C labelled protein	Assign aromatic side chains
¹³ C–NOESY-HSQC tuned to aromatic region	¹⁵ N/ ¹³ C labelled protein	Identifying short and long distance restraints from aromatic residues

The proteins were dissolved in $^2\text{H}_2\text{O}$ and a time course of ^{15}N -HSQC spectra were acquired to confirm which N-H protons exchange quickly and which remain. This was used to confirm which N-H protons may be involved in hydrogen bonding. N-H protons that are involved in H-bonding do not exchange quickly so would be present on the spectra in $^2\text{H}_2\text{O}$. N-H protons that are not involved in H-bonds and not protected from exchange would exchange with the $^2\text{H}_2\text{O}$ and therefore not be visible on the spectra. A 2D ^{15}N -NOESY-HSQC spectrum was acquired in $^2\text{H}_2\text{O}$ to identify the long and short distance restraints from aromatic amino acids.

TALOS was used to obtain a list of dihedral angle restraints. TALOS stands for Torsion Angle Likelihood Obtained from Shift and Sequence similarity. It is a database used for the prediction of Phi and Psi angles using HA, CA, CB, CO, N chemical shift assignments for a protein (Cornilescu *et al.*, 1999). TALOS creates a list of predictions for the Phi and Psi angles and rates the predictions as good, OK and bad. For structural determination all predictions rated as bad or OK were removed.

2.5.4 Structure determination using ARIA 1.2

The peaks for ^{13}C -NOESY-HSQC, ^{15}N -NOESY-HSQC, ^{13}C -HSQC tuned for the aromatic region and $^2\text{H}_2\text{O}$ -NOESY-HSQC spectra were picked manually using SPARKY (Goddard & Kneller). The spectra were transferred to Xeasy and the volumes of the peaks were determined by integration in Xeasy (Bartels *et al.*, 1995). The peak lists for the four spectra: ^{13}C -NOESY-HSQC, ^{15}N -NOESY-HSQC, ^{13}C -HSQC tuned for the aromatic region and $^2\text{H}_2\text{O}$ -NOESY-HSQC with the chemical shift lists for all the side chain atoms (determined from HCCH-TOCSY and HCONH spectra) were used for structural calculations. The list of dihedral angle restraints from TALOS was used as a restraint for structural calculations.

Structural calculations were carried out using the ARIA 1.2 software package (Linge & Nilges, 1999, Linge *et al.*, 2001, Nilges, 1995, Nilges & O' Donoghue, 1998). ARIA stands for Ambiguous Restraints for Iterative Assignments. ARIA interprets the peaks that have been picked in the NOESY spectra and defines a list of ambiguous and unambiguous NOE restraints based on the picked peaks and the chemical shift lists.

ARIA uses this information in combination with the dihedral angle restraints to produce a family of possible structures. Each ARIA run goes through 8 iterations of structures improving the structure each time and finishes by carrying out a water refinement of the 10 best energy structures it has created. At the end of the ARIA run a list of peak violations in the spectra was obtained for the determined structures. The peaks were checked manually in SPARKY to examine whether they were background noise on the spectra or real peaks. Any noise peaks would be removed.

The structures created by ARIA were displayed in the molecular graphics programme MOLMOL (Koradi *et al.*, 1996). The group of structures would be overlayed to see how similar they were. MOLMOL was used to calculate potential H-bonds in the structure. H-bonds calculated to be in 8 or more out of the 10 structures and confirmed by the presence of an N-H peak in the $^2\text{H}_2\text{O}$ - ^{15}N -HSQC spectra were included in the ARIA calculation for the subsequent run. The spectra were run through ARIA multiple times with checking of the spectra and calculating the H-bonds each time to improve the structure.

The structure was visualised in Insight II (Dayringer *et al.*, 1986) to identify the disulphide bonds. Potential disulphide bonds were added to the ARIA structure calculations as shown in table 2.5.

Table 2.5: Ambiguous disulphide bond restraints used in ARIA calculations
Cysteine 8 → Cysteine 20
Cysteine 32 → Cysteine 44
Cysteine 52 → Cysteine 65
Cysteine 59 → Cysteine 79
Cysteine 81 → Cysteine 95

For the final ARIA runs the number of structures made in iteration 8 was increased to 100 structures and the 20 best energy structures were used in the water refinement. The best energy structure for the 20 structures for wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ were used to compare the differences between MSP1₁₉ and Glu28→Lys MSP1₁₉.

The quality of the final structures was assessed using PROCHECK NMR (Laskowski *et al.*, 1996). The best energy structure for the 20 structures for wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ were used to compare the differences between MSP1₁₉ and Glu28→Lys MSP1₁₉. The structures were compared by visualisation using Insight II and by using MOLMOL (Koradi *et al.*, 1996) to calculate the surface electropotential.

Chapter 3: Antibody binding studies on individual amino acid variants

3.1 Introduction

In order to map the antibody binding sites of *P. yoelii* MSP1₁₉ a site directed mutagenesis approach was used to identify individual amino acids that may be involved in antibody binding. This approach instead of a random mutagenesis approach was used because there was information available in the literature about the potential amino acids that may be involved. The site directed mutagenesis protocol that was used is summarised in figure 3.1. Three variants with changes to residues 12, 16 and 17 were created because of the information in the literature from Benjamin *et al.* (Benjamin *et al.*, 1999) (discussed in detail in section 1.6.4) which suggested that these residues could be important for antibody binding from studies of *P. yoelii* isolates. Residue 28 was altered because it is conserved across the species and it has been shown by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) (discussed in detail in section 1.6.3) to be important for *P. falciparum* inhibitory antibody binding.

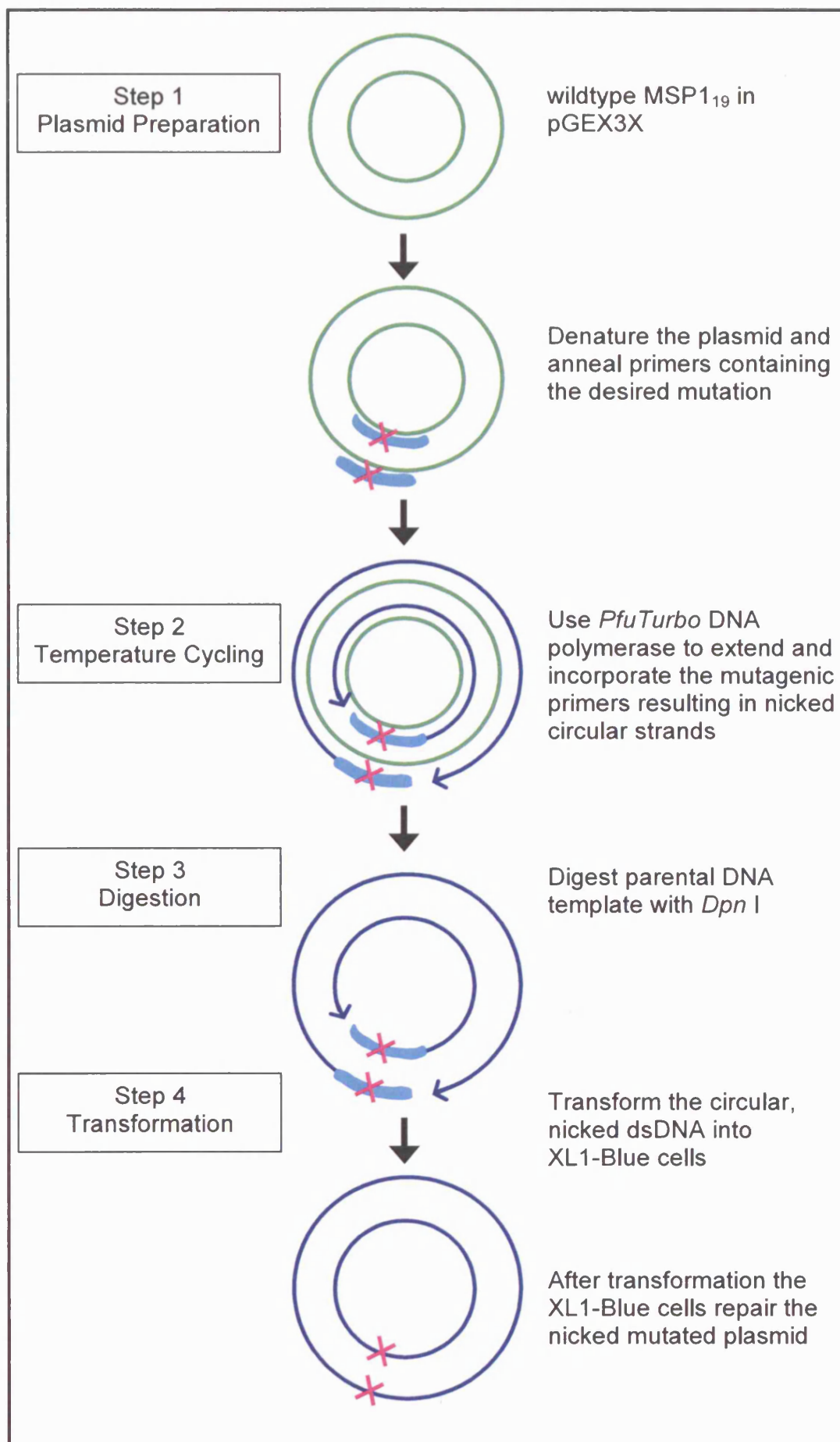
The affects of the variations on antibody binding were tested using three methods: western blotting, ELISA and surface plasmon resonance analysis. The use of three methods was to overcome any limitations of the individual methods and also to confirm the accuracy of the data obtained from those methods. Table 3.1 highlights the main advantages and disadvantages of the individual techniques.

Table 3.1: Advantages and disadvantages of western blotting, ELISA and surface plasmon resonance		
	Advantages	Disadvantages
Western blotting	Quick.	Small differences in antibody binding may not be seen due to saturation of the blot.
	Small amounts of protein can be used.	The exact amount of protein that binds to nitrocellulose is unknown.
	Cheap.	Only one antibody concentration can be tested per blot.
ELISA	A series of antibody concentrations can be set up on one plate and binding curves can be gained.	The exact amount of protein that binds to the plate is unknown.
	Quantitative antibody binding data.	There can be saturation of the signal at high antibody concentrations.
	Saturation of the signal can be overcome as a range of antibody concentrations are used.	
Surface plasmon resonance	The amount of protein bound to the chip can be accurately determined.	The chips that the proteins are bound to are very expensive.
	The amount of antibody can be accurately determined.	The binding conditions and regeneration conditions have to be optimised before analysis.

For the antibody binding studies three monoclonal antibodies against MSP1 that were created by Lilian Spencer were used called B6, F5 and B10. The antibodies and their production are discussed in detail in section 1.6.4. B6 and F5 were used because they mediate a substantial reduction in parasitaemia and all the mice injected with these antibodies clear the parasite following parasite challenge. Competition ELISA experiments showed that the epitopes for F5 and B6 overlap each other but are clearly distinct epitopes. This would suggest that different information may be gained from studying the two antibodies and areas that are important for binding both antibodies may be particularly important. B6 and F5 antibody only need the first EGF domain for antibody binding. The third antibody B10 was used because competition ELISA experiments showed that the epitope for B10 antibody was distinct from that of B6 and F5. B10 antibody caused a partial suppression of parasite growth on challenge infection. B10 antibody requires both EGF domains for binding (Spencer Valero *et al.*, 1998). Using this antibody could therefore help to determine if the variations made have an effect on the overall structure of the protein.

Figure 3.1: Overview of the site-directed mutagenesis protocol used to produce the amino acid variations in MSP1₁₉.

The QuikChange site-directed mutagenesis kit from Stratagene was used for the mutagenesis. This figure is adapted from the QuikChange site-directed mutagenesis kit manual. The parental DNA plasmid (wildtype MSP1₁₉ in pGEX3X) is shown in green. The mutagenic primers are shown in cyan and the mutated DNA plasmid is shown in blue. The mutation is indicated with the pink cross.



3.2 Expression and purification of GST-MSP1₁₉ variants

Four GST-MSP1₁₉ variants were created by site-directed mutagenesis using the sequence of wildtype MSP1₁₉ of *P. yoelii* YM in pGEX3X vector as a template. The four variants produced contained a single amino acid change as follows: Arg12→Leu; Lys16→Glu; Asn17→His and Glu28→Lys. The changes to the amino acids are shown in figure 3.2. Expression of the GST-MSP1₁₉ variants, wildtype GST-MSP1₁₉ and GST was carried out as described section 2.2. Figure 3.3 shows the expression and purification of the Asn17→His variant. In this figure there is a clear band at 37 kDa representing the GST-MSP1₁₉ in the elution fractions (lanes 6-21) from the glutathione agarose column. The expression and purification of the wildtype and other GST-MSP1₁₉ variants gave a similar purification profile (data not shown).

The purified proteins were quantified using densitometry and run on a NuPAGE gel to confirm the accuracy of the quantification (shown in figure 3.4).

3.3 Western blotting analysis of antibody binding to GST-MSP1₁₉ variants

In order to test if the amino acids changed in the GST-MSP1₁₉ variants were involved in antibody binding, western blotting analysis was carried out. 500 ng of the GST-MSP1₁₉ variants were run on NuPAGE gels under non-reducing conditions for monoclonal antibody westerns and reducing conditions for anti-GST antibody westerns. A control western blot with anti-GST antibody was carried out as shown in figure 3.5 (top panel). This western blot shows binding to all the GST-MSP1₁₉ variants confirming that the proteins have been expressed. Western blotting analysis with B6, F5 and B10 antibodies (Spencer Valero *et al.*, 1998) show differential binding with the variants (shown in figure 3.5). The results are summarised in table 3.2.

Table 3.2: Effects of amino acid variations on B6, F5 and B10 binding as shown by western blotting (figure 4)

Variant	B6	F5	B10
Arg12→Leu	++	+	++
Lys16→Glu	-	-	++
Asn17→His	+	++	++
Glu28→Lys	++	-	++
++ → binding equivalent to wildtype binding			
+ → reduced binding compared to wildtype binding			
- → no binding			

3.4 ELISA analysis of antibody binding to GST-MSP1₁₉ variants

ELISA was used to analyse the kinetics of antibody binding to GST-MSP1₁₉ variants over a range of antibody concentrations. Initial ELISA experiments involved using a range of concentrations of goat anti-GST antibody to determine the optimum concentration for wildtype GST-MSP1₁₉ capture. The results are shown in figure 3.6, this shows an increase in ELISA signal up to 3 µg/ml and indicates an optimum concentration of 0.4 µg/ml.

1 µg/ml GST-MSP1₁₉ variants were bound to 0.4 µg/ml goat-anti-GST antibody bound to the ELISA plate. The proteins were probed with doubling dilutions of B6, F5 or B10 antibody and 1/2000 dilution anti-mouse IgG HRP conjugate. The peroxidase was detected and absorbance was read at 490 nm. The ELISA shows clear differences in the binding curves of the GST-MSP1₁₉ variants (Figures 3.7, 3.8, 3.9). The differences are summarised in table 3.3.

Table 3.3: Effects of amino acid variations on B6, F5 and B10 binding curves in ELISA experiments (Figures 3.7, 3.8, 3.9)

Variant	B6	F5	B10
Arg12→Leu	++	+	+
Lys16→Glu	-	-	+
Asn17→His	+	++	+
Glu28→Lys	+	-	+
++ → binding equivalent to wildtype binding			
+ → reduction in the steepness of the binding curve			
- → no binding			

The results of the F5 ELISA with all the variants are in agreement with the western blotting analysis. The results of the B6 ELISA experiments for the Glu28→Lys do not agree with the western blotting results since the ELISA shows a reduction in binding whereas the western blotting shows no change in binding. The B10 ELISA shows a reduction in binding for all the variants whereas the western blotting shows no change in binding.

3.5 SPR of antibody binding to GST-MSP1₁₉ variants

SPR analysis allows protein interactions to be detected in real time. The technique is explained in the schematic in figure 3.10. A sensor chip that is coated in a thin layer of gold is used. The protein is immobilised onto the surface of the chip. The ligand is passed across the surface of the chip in solution in the flow cell. Polarised light is shone at the sensor chip and reflected off. When buffer is passed over the chip with no ligand to bind the reflected light may be at an angle shown for point 1 (in figure 3.10). This would form the base line of the sensorgram. When the ligand solution flows across the chip, ligand would bind to the protein on the chip surface. This alters the refractive index at the interface between the chip surface and the ligand solution. This reduces the angle of reflected light. This change in angle is proportional to the mass of the bound material and is recorded on the sensorgram in arbitrary units called resonance units. For

example, in figure 3.10, when a small amount is bound as shown at point 2 the angle of reflected light is reduced and the resonance units on the sensorgram go up. This is because the resonance units are inversely proportional to the angle of the reflected light. As more ligand binds as shown at point 3, the signal goes up further. After binding as the buffer flows over, the ligand dissociates. To remove all of the ligand a buffer with low or high pH is used to disrupt the protein-ligand binding interactions.

For this project, the GST-MSP1₁₉ variants were bound to anti-GST antibody immobilised on the surface of a CM5 chip and binding assays were performed with B6, F5 and B10 antibodies. A schematic of the sensorgram for the SPR experiments is shown in figure 3.11 indicating the positions where the binding levels are recorded. The binding in resonance units was converted into percentages to allow comparison between the variants (as described in section 2.2.6). The results show the percentage binding, in resonance units, of the antibody compared between the wildtype and variants (figures 3.12, 3.13, 3.14). . The SPR analysis shows clear differences in antibody binding for the GST-MSP1₁₉ variants. The differences are summarised in table 3.4.

Table 3.4: Effects of amino acid variations on B6, F5 and B10 antibody binding detected by SPR analysis (Figures 3.12, 3.13, 3.14)			
Variant	B6	F5	B10
Arg12→Leu	++	+	+
Lys16→Glu	-	-	++
Asn17→His	+	++	++
Glu28→Lys	+	-	-
++ → binding equivalent to wildtype binding			
+ → small reduction in antibody binding			
- → large reduction in antibody binding			

The results for B6 and F5 binding are in agreement with the ELISA data. The results for B10 binding in the SPR analysis show a small reduction for the Arg12→Leu variant and a significant reduction of 45 percent for the Glu28→Lys variant which was not picked up by the western blotting.

Figure 3.2: Single amino acid variations made to wildtype MSP1₁₉.

The left hand panel shows the structures of the amino acids in the wildtype protein. The right hand panel shows the structures the amino acids that have been substituted in the variants. The areas of the amino acid structures that are different are highlighted in pink. This figure was produced using MDL ISIS/Draw 2.5.

Wildtype	Variant
$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{N}-\text{C}-\text{NH}_2 \\ \quad \quad \quad \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{H}_2 \quad \text{H}_2 \quad \text{H} \quad \text{NH}_2^+ \end{array} $ <p>Arginine 12</p>	$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{CH}(\text{CH}_3)_2 \\ \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{CH}_3 \end{array} $ <p>Leucine 12</p>
$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{C}-\text{NH}_3^+ \\ \quad \quad \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{H}_2 \quad \text{H}_2 \quad \text{H}_2 \end{array} $ <p>Lysine 16</p>	$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}(=\text{O})\text{O}^- \\ \quad \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{H}_2 \quad \text{O} \end{array} $ <p>Glutamic acid 16</p>
$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}(=\text{O})\text{NH}_2 \\ \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{O} \end{array} $ <p>Asparagine 17</p>	$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}_2\text{H}_3\text{N}_3 \\ \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{O} \end{array} $ <p>Histidine 17</p>
$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}(=\text{O})\text{O}^- \\ \quad \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{H}_2 \quad \text{O} \end{array} $ <p>Glutamic acid 28</p>	$ \begin{array}{c} \\ \text{H}-\text{N} \\ \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{C}-\text{NH}_3^+ \\ \quad \quad \quad \quad \\ \text{O}=\text{C} \quad \text{H}_2 \quad \text{H}_2 \quad \text{H}_2 \quad \text{H}_2 \end{array} $ <p>Lysine 28</p>

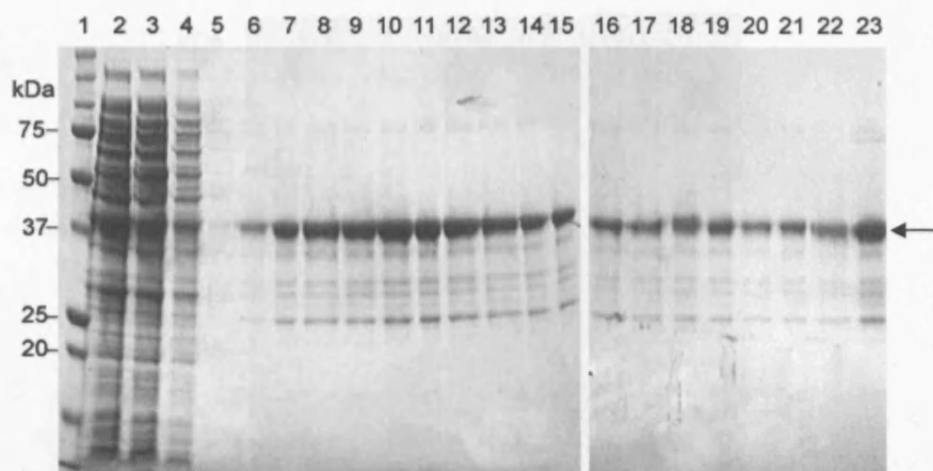


Figure 3.3: NuPAGE gel analysis of Asn17→His GST-MSP₁₉ variant expression and purification.

Asn17→His GST-MSP₁₉ variant was expressed in a 610 ml bacterial culture by IPTG induction (1 mM) for 3 hours. The protein was purified using a glutathione agarose column eluting in 0.5 ml fractions with 5 mM reduced glutathione. The protein samples were run on a pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gel in MOPS buffer under reducing conditions and stained with Coomassie blue. Lane 1 – molecular mass markers, lane 2 – cell lysate before purification, lane 3 – flow through from glutathione agarose column, lanes 4, 5 – column washes, lane 6-21 – elution fractions, lanes 22, 23 – elution fractions pooled and dialysed against PBS. The band indicated by the arrow at 37 kDa is the Asn17→His GST-MSP₁₉ variant.

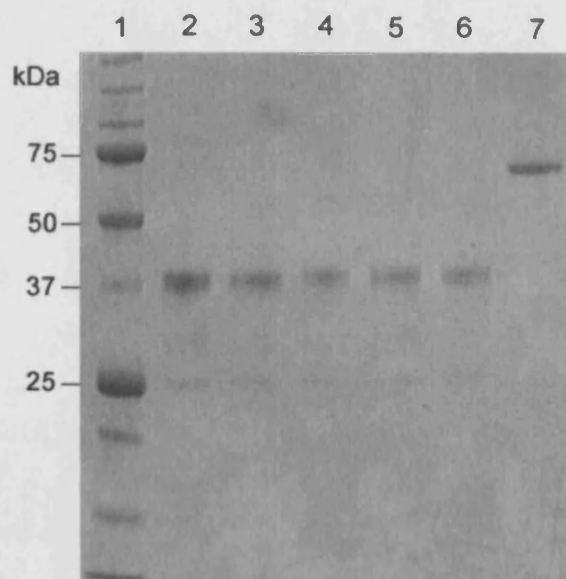
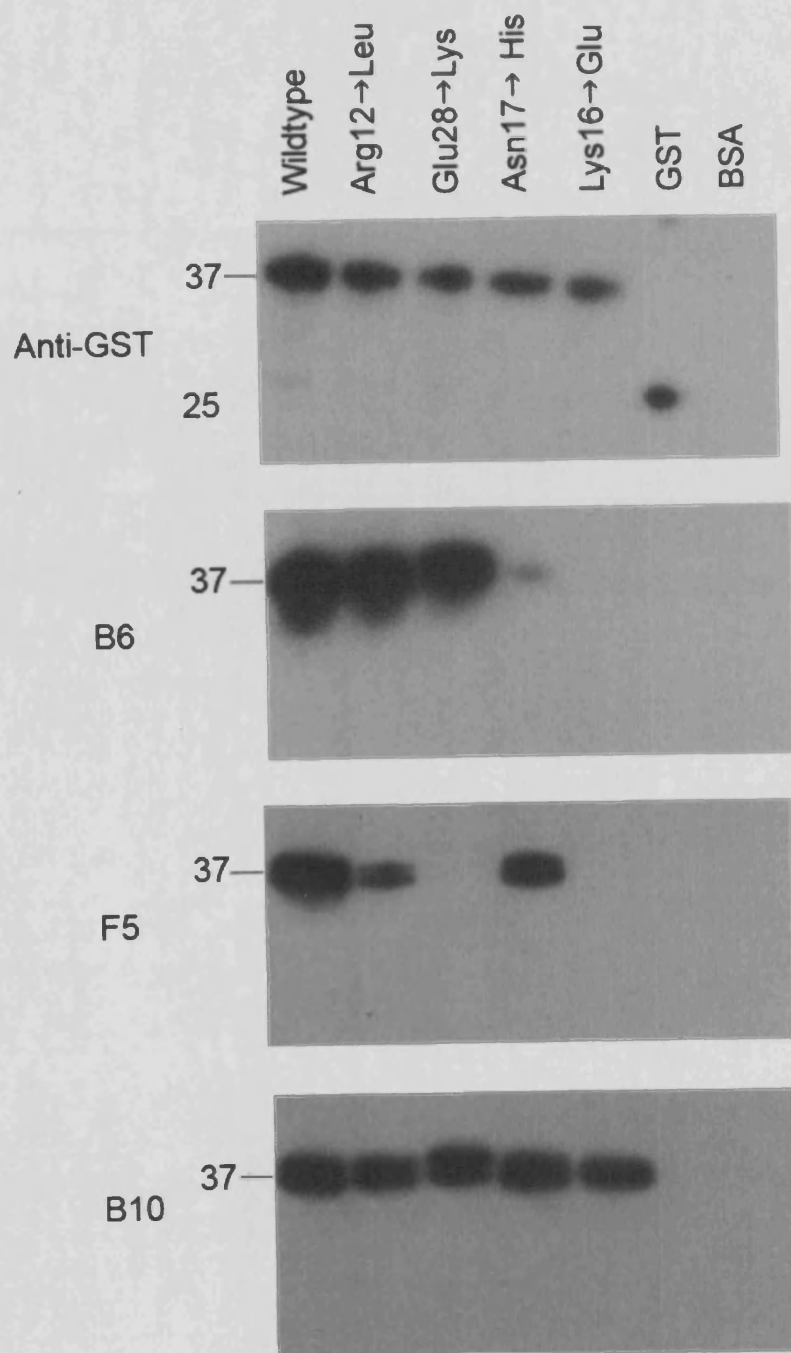


Figure 3.4: NuPAGE gel analysis of 500 ng of GST-MSP₁₁₉ variants quantified by densitometry.

500 ng of the GST-MSP₁₁₉ variants was run on a pre-cast 12 % NuPAGE Bis-Tris polyacrylamide gel in MOPS buffer under reducing conditions and visualised by Coomassie blue staining. Lane 1 – molecular mass markers, lane 2 – wildtype GST-MSP₁₁₉, lane 3 – Arg12→Leu, lane 4 – Glu28→Lys, lane 5 – Asn17→His, lane 6 – Lys16→Glu, lane 7 – 500 ng purified BSA.

Figure 3.5: Western blotting analysis of antibody binding to GST-MSP1₁₉ variants.

500 ng of the wildtype GST-MSP1₁₉ and GST-MSP1₁₉ variants were run on NuPAGE gels under reducing conditions (anti-GST antibody western blots) or non-reducing conditions (B6, F5, B10 antibody western blots) and transferred to nitrocellulose. GST and BSA were used as controls. The anti-GST western blot (top panel) was probed with 1/5000 dilution anti-GST HRP conjugate. The B6, F5 and B10 antibody western blots were first probed with B6 (2 µg/ml), F5 (10 µg/ml) or B10 (2 µg/ml) followed by 1/2000 dilution goat anti-mouse IgG HRP conjugate. The bands at 37 kDa are the GST-MSP1₁₉ proteins.



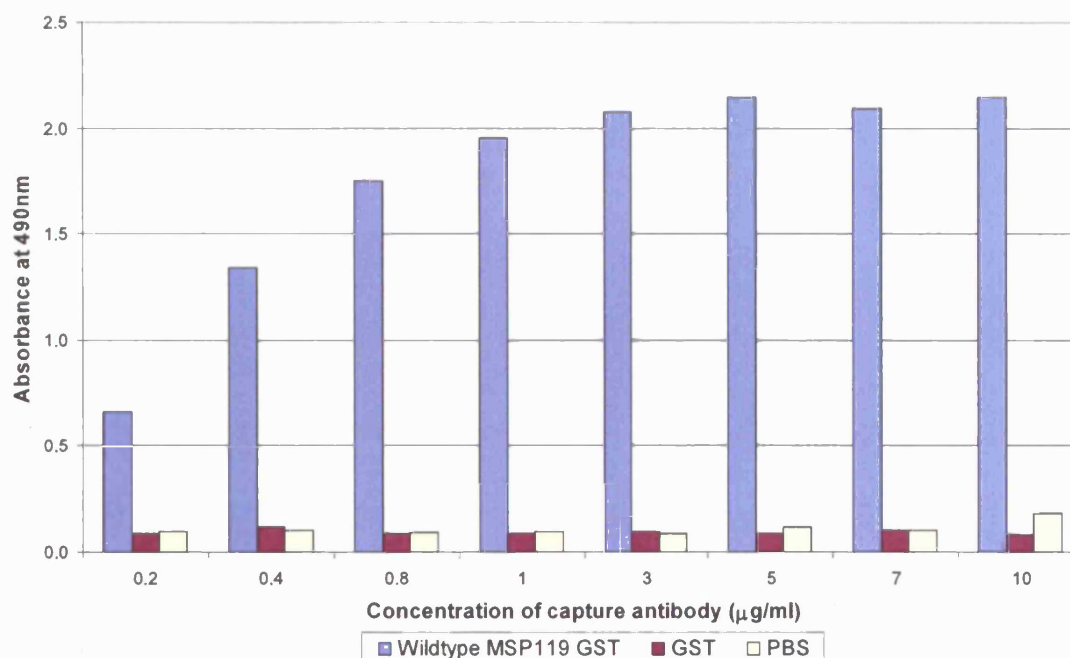


Figure 3.6: Antibody sandwich ELISA to determine the optimum concentration of capture antibody.

Goat anti-GST antibody at a range of concentrations between 0.2 and 10 $\mu\text{g/ml}$ was used to coat the ELISA plate as the capture antibody. 1 $\mu\text{g/ml}$ of wildtype GST-MSP1₁₉, GST or PBS was bound to the capture antibody. This was probed with 1.6125 $\mu\text{g/ml}$ B6 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. Duplicate plates were used and the mean results are shown on the graph.

Figure 3.7: ELISA of B6 antibody binding to GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B6 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Lys16→Glu (KE16) is shown in green, Asn17→His (NH17) is shown in yellow, Arg12→Leu (RL12) is shown in blue and Glu28→Lys (EK28) is shown in pink.

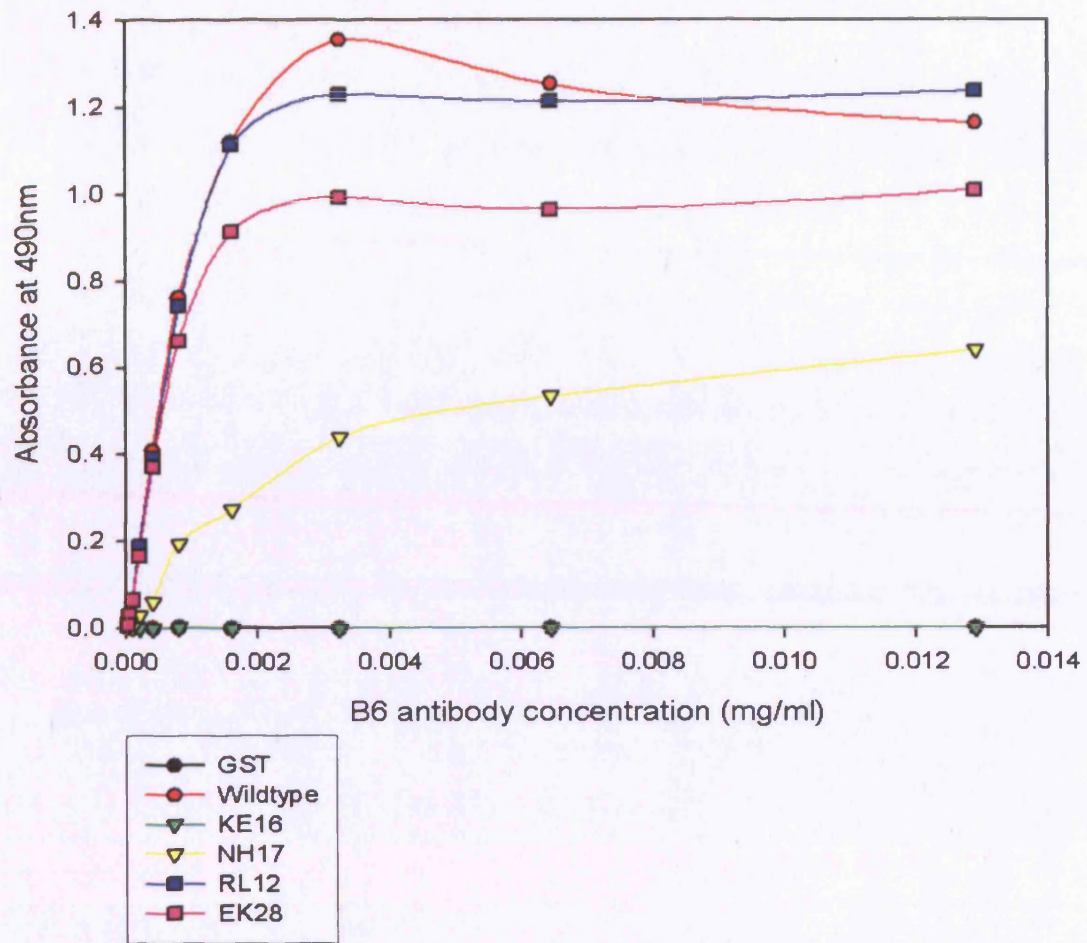


Figure 3.8: ELISA of F5 antibody binding to GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of F5 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Lys16→Glu (KE16) is shown in green, Asn17→His (NH17) is shown in yellow, Arg12→Leu (RL12) is shown in blue and Glu28→Lys (EK28) is shown in pink.

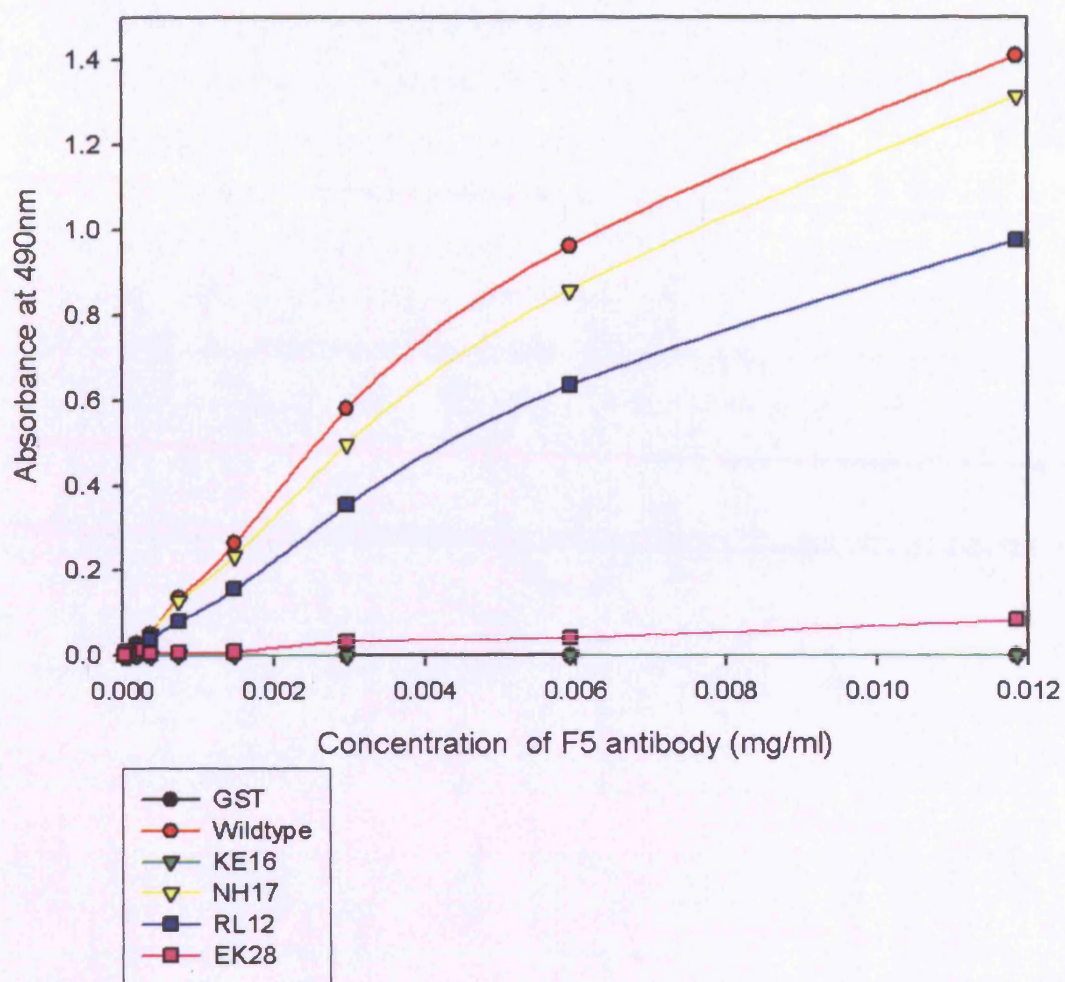


Figure 3.9: ELISA of B10 antibody binding to GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B10 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Lys16→Glu (KE16) is shown in green, Asn17→His (NH17) is shown in yellow, Arg12→Leu (RL12) is shown in blue and Glu28→Lys (EK28) is shown in pink.

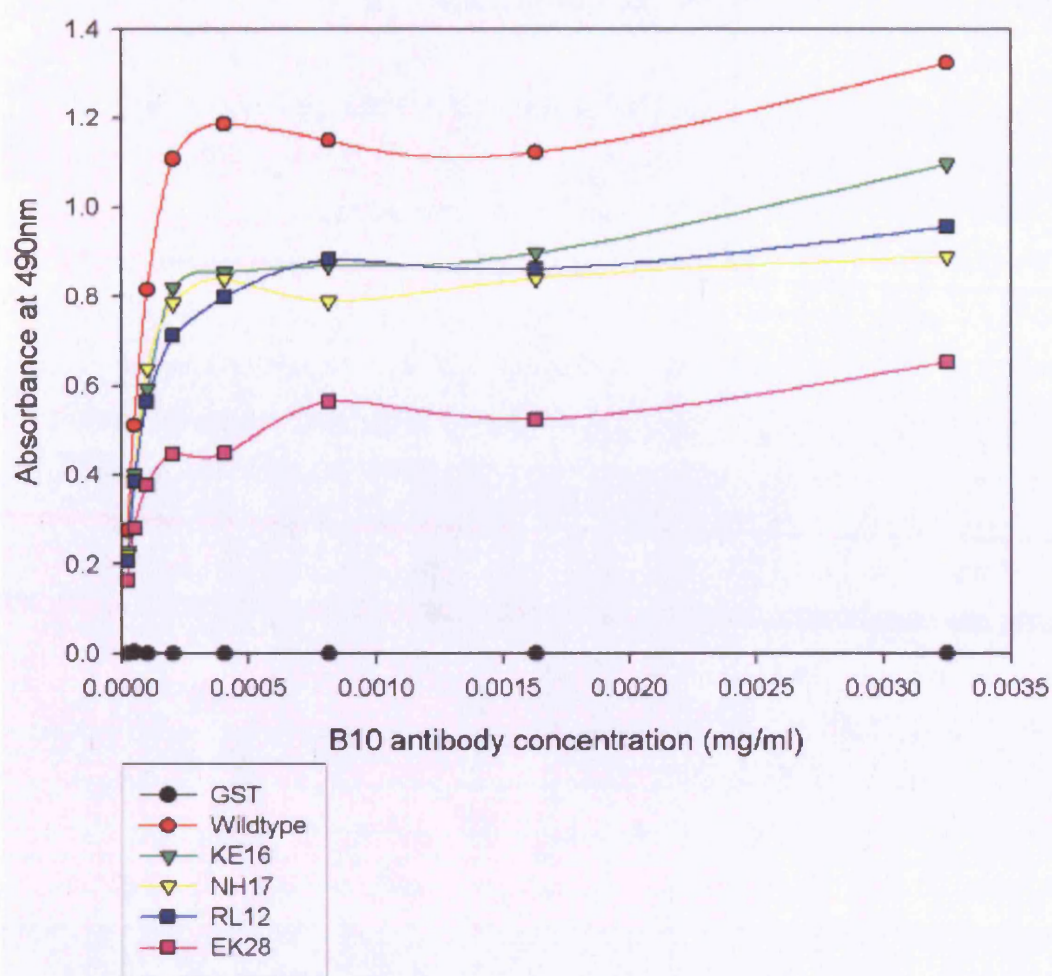


Figure 3.10: Overview of surface plasmon resonance.

A: This panel shows a schematic of the surface plasmon resonance equipment.

B: This panel shows a schematic of a typical sensorgram that is obtained.

The protein is immobilised onto the surface of the chip. The ligand is passed across the surface of the chip in solution in the flow cell. Polarised light is shone at the sensor chip and reflected off. When buffer is passed over the chip with no ligand to bind the reflected light may be at an angle shown for point 1. This would form the base line of the sensorgram (B point 1). When the ligand solution flows across the chip, ligand would bind to the protein on the chip surface. This alters the refractive index at the interface between the chip surface and the ligand solution. This reduces the angle of reflected light. This change in angle is proportional to the mass of the bound material and is recorded on the sensorgram in arbitrary units called resonance units. For example, when a small amount is bound as shown at point 2 the angle of reflected light is reduced and the resonance units on the sensorgram go up (B point 2). This is because the resonance units are inversely proportional to the angle of the reflected light. As more ligand binds as shown at point 3, the signal goes up further. After binding as the buffer flows over, the ligand dissociates.

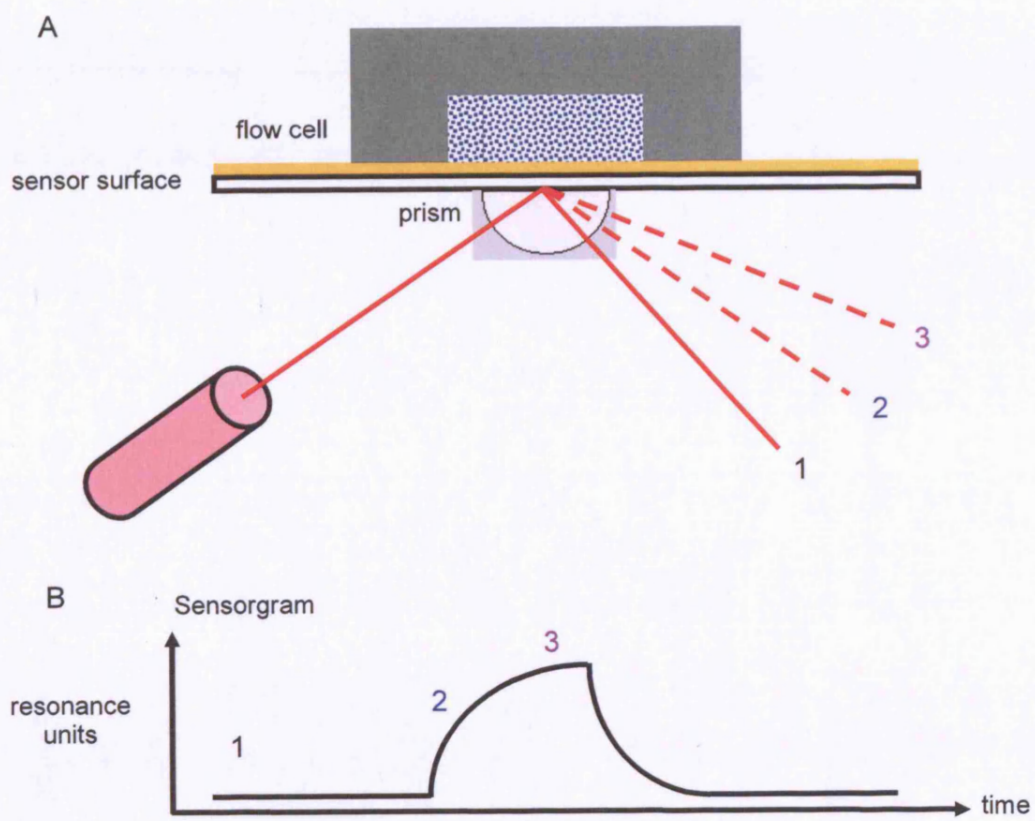
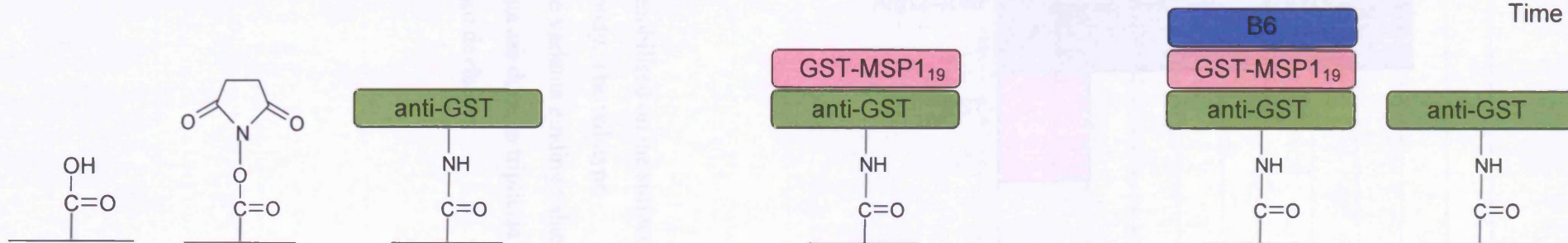
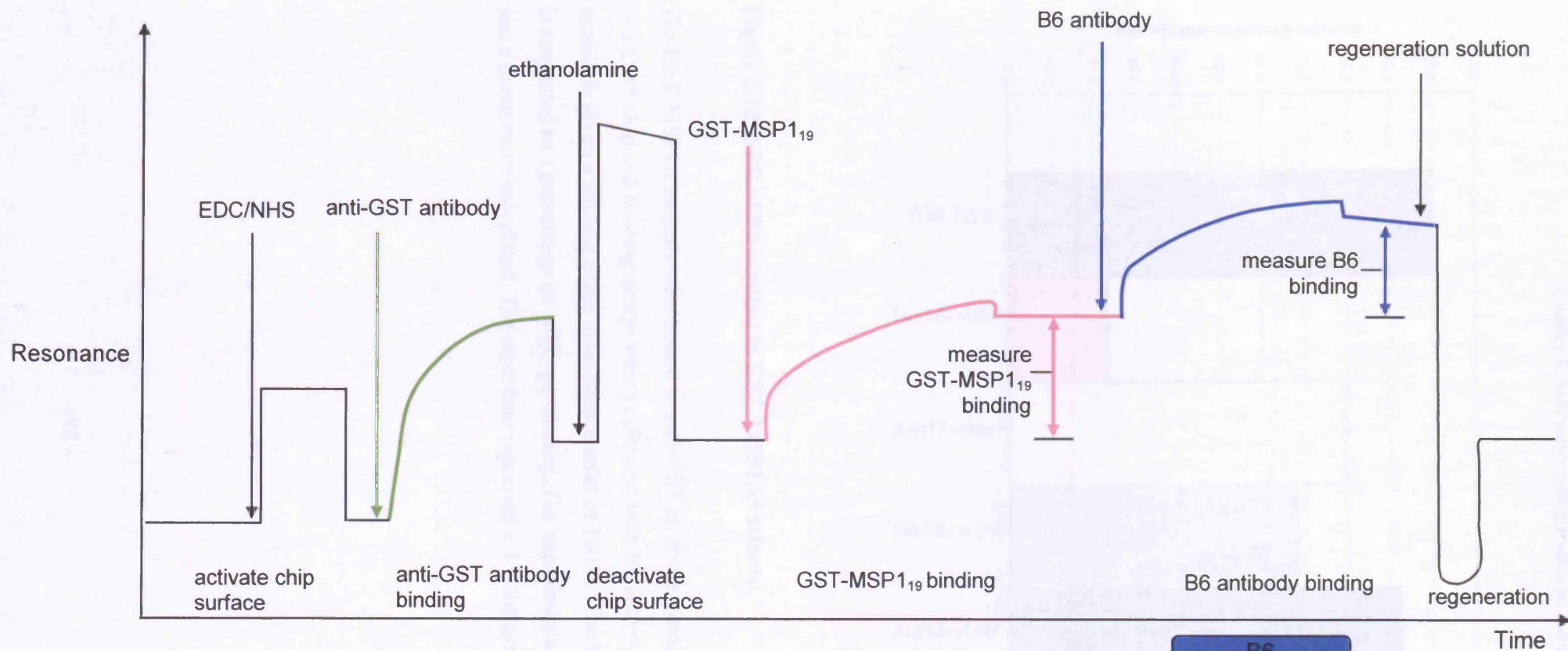


Figure 3.11: Schematic of surface plasmon resonance sensorgram obtained in the experiments in this project.

The top panel shows a schematic representation of the sensorgram. The arrows indicate the injections of solutions over the chip. The injection and binding of anti-GST antibody is highlighted in green. The injection and binding of GST-MSP1₁₉ is highlighted in pink. The injection and binding of B6 antibody is highlighted in blue. The measurements of GST-MSP1₁₉ binding and B6 antibody binding are indicated with double headed arrows.

The bottom panel shows the appearance of the chip surface following the injections of solutions over the chip.



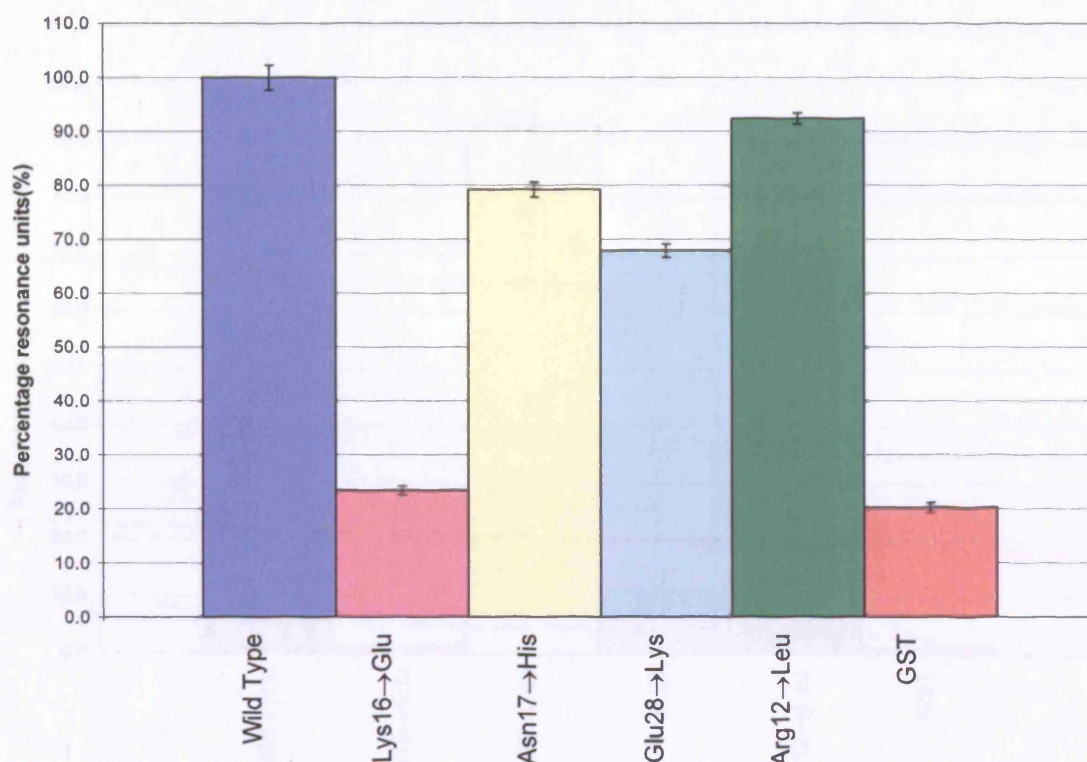


Figure 3.12: SPR of B6 binding to GST-MSP1₁₉ variants.

The GST-MSP1₁₉ variants were bound to anti-GST antibody immobilised on the surface of a CM5 chip and binding assays were performed with B6 antibody. The wildtype protein is given a binding value in resonance units of 100 %. The variants binding value is converted to a percentage of wildtype binding. The experiments are done in triplicate and a mean value calculated. The error bars represent ± 1 standard deviation.

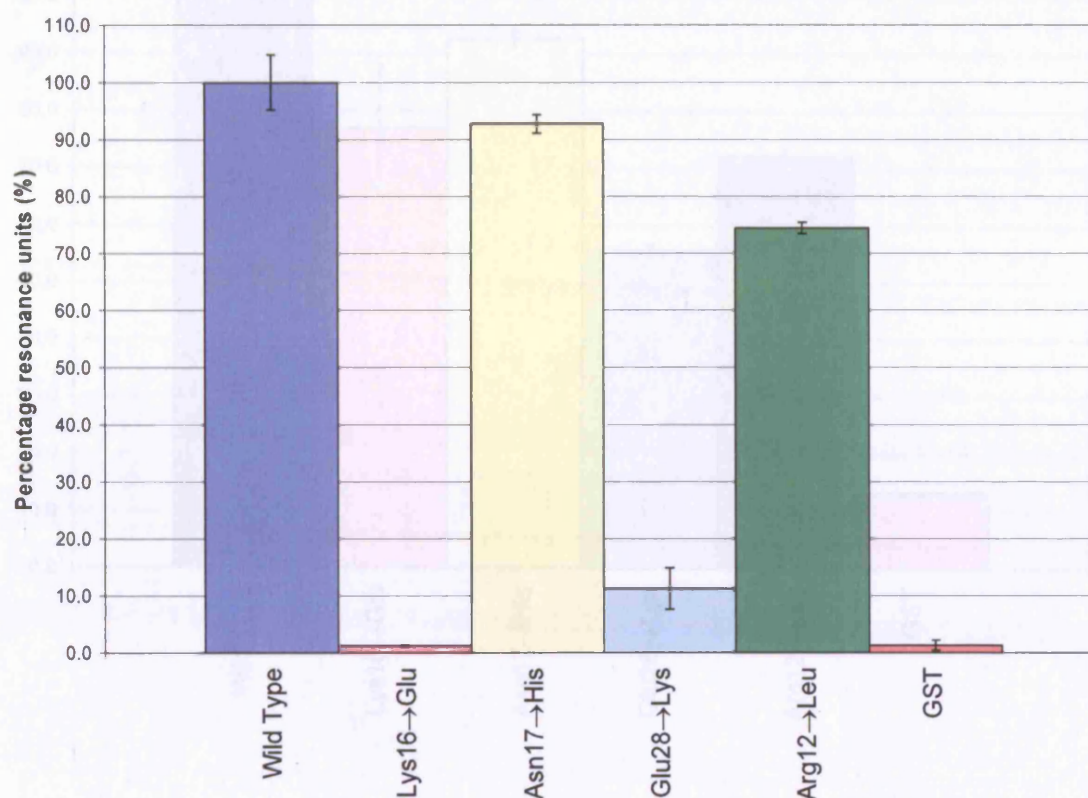


Figure 3.13: SPR of F5 binding to GST-MSP1₁₉ variants.

The GST-MSP1₁₉ variants were bound to anti-GST antibody immobilised on the surface of a CM5 chip and binding assays were performed with F5 antibody. The wildtype protein is given a binding value in resonance units of 100 %. The variants binding value is converted to a percentage of wildtype binding. The experiments are done in triplicate and a mean value calculated. The error bars represent ± 1 standard deviation.

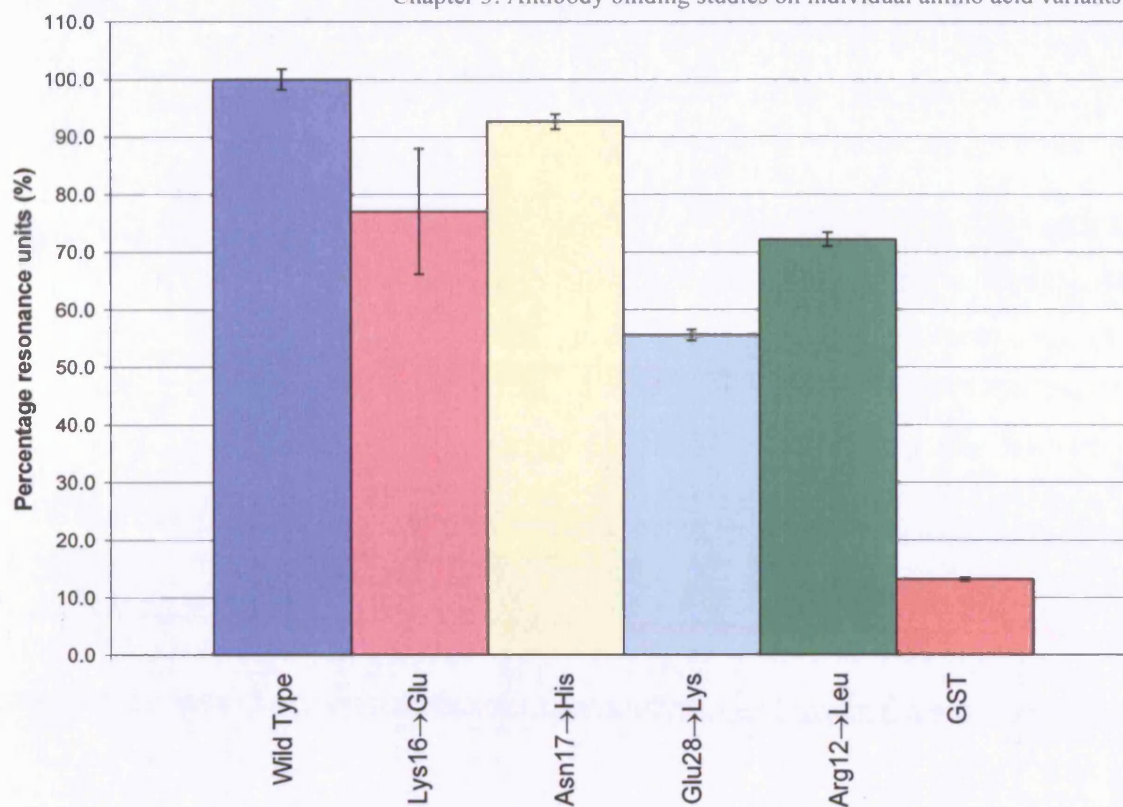
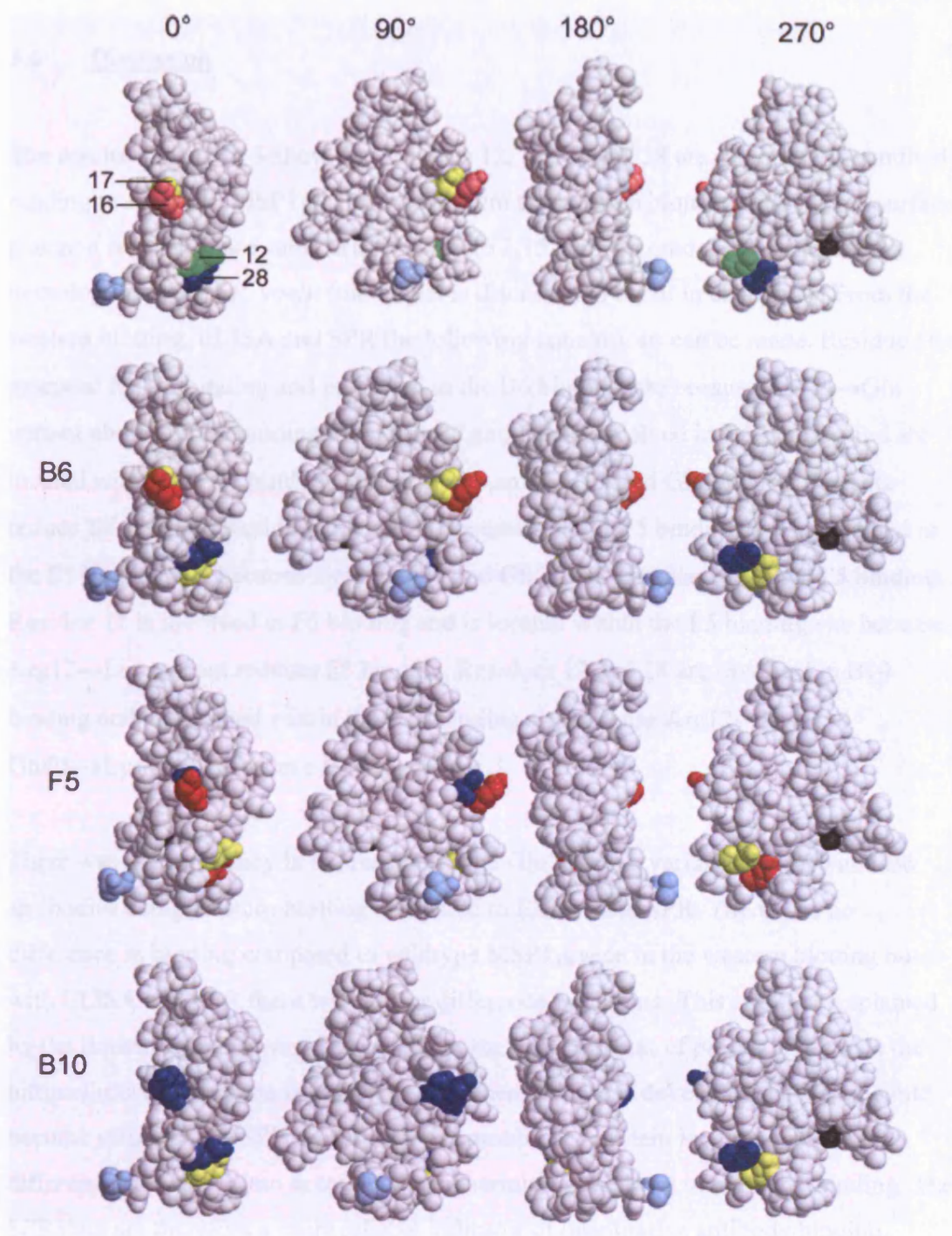


Figure 3.14: SPR of B10 binding to GST-MSP1₁₉ variants.

The GST-MSP1₁₉ variants were bound to anti-GST antibody immobilised on the surface of a CM5 chip and binding assays were performed with B10 antibody. The wildtype protein is given a binding value in resonance units of 100 %. The variants binding value is converted to a percentage of wildtype binding. The experiments are done in triplicate and a mean value calculated. The error bars represent ± 1 standard deviation.

Figure 3.15: Summary of the effect of single amino acid changes on B6, F5 and B10 antibody binding to GST-MSP1₁₉.

A homology model of *P. yoelii* MSP1₁₉ was created using the Swiss Model Protein Modelling Server (top panel). This figure highlights the 3-D location of the amino acids changed in the variant proteins. Arg12→Leu is shown in green, Lys16→Glu is shown in pink, Asn17→His is shown in yellow and Glu28→Lys is shown in blue. The C-terminal residue is shown in black and the N-terminal residue is shown in light blue. The effects of the amino acid changes on B6, F5 and B10 antibody binding observed in the western blotting, ELISA and SPR experiments are highlighted. The residues that have no effect on binding are shown in blue, those that result in partial reduction in binding are shown in yellow and those that abolish binding are shown in red.



3.6 Discussion

The results of chapter 3 show that residues 12, 16, 17 and 28 are important for antibody binding to *P. yoelii* MSP1₁₉. The results from the western blotting, ELISA and surface plasmon resonance are summarised in figure 3.15 mapped onto the 3-Dimensional homology model of *P. yoelii* (the model is discussed in detail in chapter 6). From the western blotting, ELISA and SPR the following conclusions can be made. Residue 16 is essential for B6 binding and is located in the B6 binding site because Lys16→Glu variant abolishes B6 binding. Residues 17 and 28 are involved in B6 binding and are located within the B6 binding site because Asn17→His and Glu28→Lys variants reduce B6 binding. Residues 16 and 28 are essential for F5 binding and are located in the F5 binding site because Lys16→Glu and Glu28→Lys variants abolish F5 binding. Residue 12 is involved in F5 binding and is located within the F5 binding site because Arg12→Leu variant reduces F5 binding. Residues 12 and 28 are involved in B10 binding and are located within the B10 binding site because Arg12→Leu and Glu28→Lys variants reduce B10 binding.

There was a discrepancy in the results for the Glu28→Lys variant with B10 and B6 antibodies using western blotting compared to ELISA and SPR. There was no difference in binding compared to wildtype MSP1₁₉ seen in the western blotting but with ELISA and SPR there was a clear difference in binding. This could be explained by the limitations with western blotting as the exact amount of protein that binds the nitrocellulose membrane is unknown and when the blot is developed the signal could become saturated. In SPR the amount of immobilised protein is measured and the differences are taken into account when determining the level of antibody binding. The SPR data are therefore a more reliable indicator of quantitative antibody binding.

Spencer *et al.* carried out competitive ELISA analysis using antibodies against *P. yoelii* MSP1₁₉. The analysis indicated that the epitopes for B6 and F5 antibodies overlapped each other but were clearly distinct (Spencer Valero *et al.*, 1998). The data in this project supports these findings as they indicate that changes to some residues, e.g. residue 16, abolish binding to both antibodies and are hence likely to be located in both

binding sites and changes to some residues e.g. residue 12, only affect binding to one of the antibodies. The data in this project suggests that the B10 antibody binding epitope overlaps with the F5 and B6 antibody binding epitopes because changing residue 28 affects F5, B6 and B10 antibody binding suggesting it is in all three binding sites. This disagrees with the competition ELISA from Spencer *et al.* (Spencer Valero *et al.*, 1998) which suggests that the B10 epitope does not overlap with F5 or B6. The reason for the difference may be because of the methods used. In competitive ELISA analysis there may need to be significant overlap between the B10 and B6 or F5 antibody binding sites in order to see an effect on binding whereas by site directed mutagenesis individual residues in a binding site can be identified.

Benjamin *et al.* studied antibody binding to sequence variants from *P. yoelii* isolates expressed in bacteria. They indicated Lys16→Glu and Asn17→His in *P. yoelii* isolates correlated with affecting binding to B6 and F5 antibodies (Benjamin *et al.*, 1999). The data in this project for the Lys16→Glu variant agrees with the findings of Benjamin *et al.* (Benjamin *et al.*, 1999). The data in this project indicate that Asn17→His does not affect F5 antibody binding and only reduces B6 antibody binding. This difference may be explained as the isolates containing the Asn17→His variation also contained multiple sequence variations including the Lys16→Glu variation and therefore the effect of Asn17→His alone could not be clearly identified. Benjamin *et al.* indicated that Arg12→Ser in isolates correlated with no binding to B6, F5 and B10 antibodies (Benjamin *et al.*, 1999). In this report an Arg12→Leu variation was made rather than an Arg12→Ser variation because arginine and leucine are similar in size. These data indicated that Arg12→Leu was important for reduction in F5 and B10 antibody binding only, suggesting this is not as important for antibody binding as suggested by Benjamin *et al.* (Benjamin *et al.*, 1999).

Uthaipibull *et al.* produced a Glu28 variant in *P. falciparum* which was found to effect inhibitory antibody binding (Uthaipibull *et al.*, 2001). This was in agreement with the *P. yoelii* data presented in this project for the Glu28→Lys variant which shows that Glu28→Lys is important for antibody binding in *P. yoelii*. This implies that residue 28 is important for antibody binding across the species and its conservation may be of functional importance.

Chapter 4: Designing a double MSP1₁₉ variant to affect all three monoclonal antibodies

4.1 Introduction

In this chapter, I will discuss the design of a double MSP1₁₉ variant based on the four individual amino acid variants discussed in chapter 3 with the aim of affecting binding to B6, F5 and B10 monoclonal antibodies. The Lys16→Glu MSP1₁₉ variant abolished binding to B6 and F5 antibody. This variant therefore provided a good starting point for producing a double variant that could affect binding to all three monoclonal antibodies. The Glu28→Lys MSP1₁₉ variant was the only one of the four variants that significantly reduced binding to B10 antibody. A double variant was therefore created with two amino acid changes as follows: Lys16→Glu and Glu28→Lys. The hypothesis for antibody binding to the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant would be that it would have the combined effect of the two individual variants. The hypothesis would therefore be that the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant would not bind to F5 and B6 and there would be a reduction in binding to B10. In this chapter, I will compare the binding of the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant to the individual Lys16→Glu and Glu28→Lys GST-MSP1₁₉ variants.

4.2 Expression and purification of double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant

A double Lys16→Glu /Glu 28→Lys GST-MSP1₁₉ variant was created by site-directed mutagenesis using the Lys16→Glu MSP1₁₉ gene of *P. yoelii* YM in pGEX3X vector as a template. Expression of the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant was carried out as described section 2.2. The protein was purified using a glutathione agarose column and quantified using densitometry. Figure 4.1 shows the purification of the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant.

4.3 Western blotting analysis of antibody binding to residues 16 and 28 single and double GST-MSP1₁₉ variants

Western blotting analysis was carried out to compare the binding of the double Lys16→Glu /Glu28→Lys GST-MSP1₁₉ variant to the individual Lys16→Glu and Glu28→Lys GST-MSP1₁₉ variants. 500 ng of the wildtype GST-MSP1₁₉, double Lys16→Glu /Glu28→Lys, Lys16→Glu and Glu28→Lys GST-MSP1₁₉ variants were run on NuPAGE gels under non-reducing conditions for monoclonal antibody westerns and reducing conditions for anti-GST antibody westerns. A control western blot with anti-GST antibody was carried out as shown in figure 4.2 (top panel). This western blot shows binding to the wildtype, individual and double GST-MSP1₁₉ variants confirming that the double Lys16→Glu/Glu28→Lys variant has been purified and quantified in the same way as the previously produced individual GST-MSP1₁₉ variants. Western blotting analysis with B6, F5 and B10 antibodies (Spencer Valero *et al.*, 1998) show that the effect of the double Lys16→Glu/Glu28→Lys variant was a combination of the effects of the individual residue 16 and 28 variants. The results are summarised in table 4.1.

Table 4.1: Effects of double and single amino acid variations on B6, F5 and B10 binding as shown by western blotting (figure 4.2)			
Variant	B6	F5	B10
Lys16→Glu	-	-	++
Glu28→Lys	++	-	++
double Lys16→Glu/Glu28→Lys	-	-	++
++ → binding equivalent to wildtype binding			
- → no binding			

4.4 **ELISA analysis of antibody binding to residues 16 and 28 double and single GST-MSP1₁₉ variants**

ELISA was used to analyse the kinetics of antibody binding to the double and single GST-MSP1₁₉ variants over a range of antibody concentrations to look for smaller differences in binding that may not have been seen on the western blotting. The ELISA experiments were carried out using the ELISA conditions that were optimised for the analysis of the single amino acid variants described in chapter 3.

1 µg/ml wildtype, double Lys16→Glu/Glu28→Lys, Lys16→Glu and Glu28→Lys GST-MSP1₁₉ variants were bound to 0.4 µg/ml goat-anti-GST antibody bound to the ELISA plate. The proteins were probed with doubling dilutions of B6, F5 or B10 antibody and 1/2000 dilution anti-mouse IgG HRP conjugate. The peroxidase was detected and absorbance was read at 490 nm. The ELISA results showed that the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant gave binding curves that showed a combination of the effects of the two single variants. The binding curve for B10 antibody showed an additive effect of the reduction in steepness in binding curve seen for the two single variants with the double variant giving a larger reduction in the steepness of the binding curve than the single variants. The results are summarised in table 4.2.

Table 4.2: Effects of single and double amino acid variations on B6, F5 and B10 binding curves in ELISA experiments (Figures 4.3, 4.4, 4.5)

Variant	B6	F5	B10
Lys16→Glu	-	-	++
Glu28→Lys	+	-	+
double Lys16→Glu/Glu28→Lys	-	-	+
++ → small reduction in the steepness of the binding curve			
+ → large reduction in the steepness of the binding curve			
- → no binding			

Figure 4.1: NuPAGE gel analysis of double Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant expression and purification.

Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant was expressed in a 610 ml bacterial culture by IPTG induction (1 mM) for 3 hours. The protein was purified using a glutathione agarose column eluting in 0.5 ml fractions with 5 mM reduced glutathione. The protein samples were run on a pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gel in MOPS buffer under reducing conditions and stained with Coomassie blue. Lane 1 – molecular mass markers, lane 2 – cell lysate before purification, lane 3 – flow through from glutathione agarose column, lanes 4-5 – column washes, lane 6-7 – elution fractions pooled and dialysed against PBS. The band indicated by the arrow at 37 kDa is the Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant.

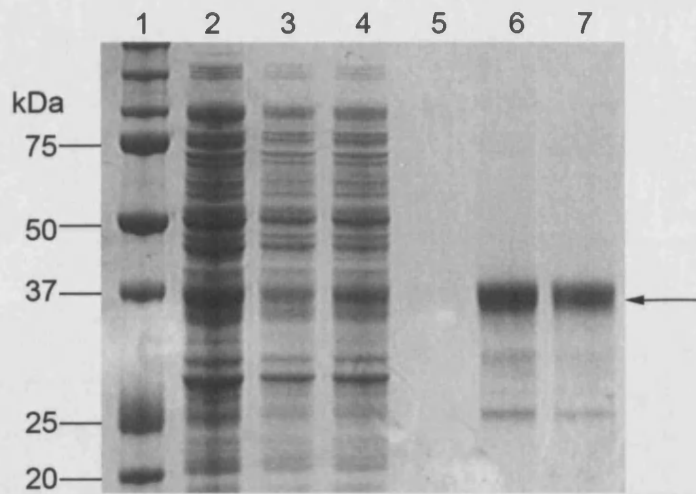


Figure 4.2: Western blotting analysis of antibody binding to residues 16 and 28 single and double GST-MSP₁₁₉ variants.

500 ng of the wildtype, Lys16→Glu, Glu28→Lys and Lys16→Glu/Glu28→Lys GST-MSP₁₁₉ variants were run on NuPAGE gels under reducing conditions (anti-GST antibody western blots) or non-reducing conditions (B6, F5, B10 antibody western blots) and transferred to nitrocellulose. GST was used as a control. The anti-GST western blot (top panel) was probed with 1/5000 dilution anti-GST HRP conjugate. The B6, F5 and B10 antibody western blots were first probed with B6 (2 µg/ml), F5 (10 µg/ml) or B10 (2 µg/ml) followed by 1/2000 dilution goat anti-mouse IgG HRP conjugate. The bands at 37 kDa are the GST-MSP₁₁₉ proteins.

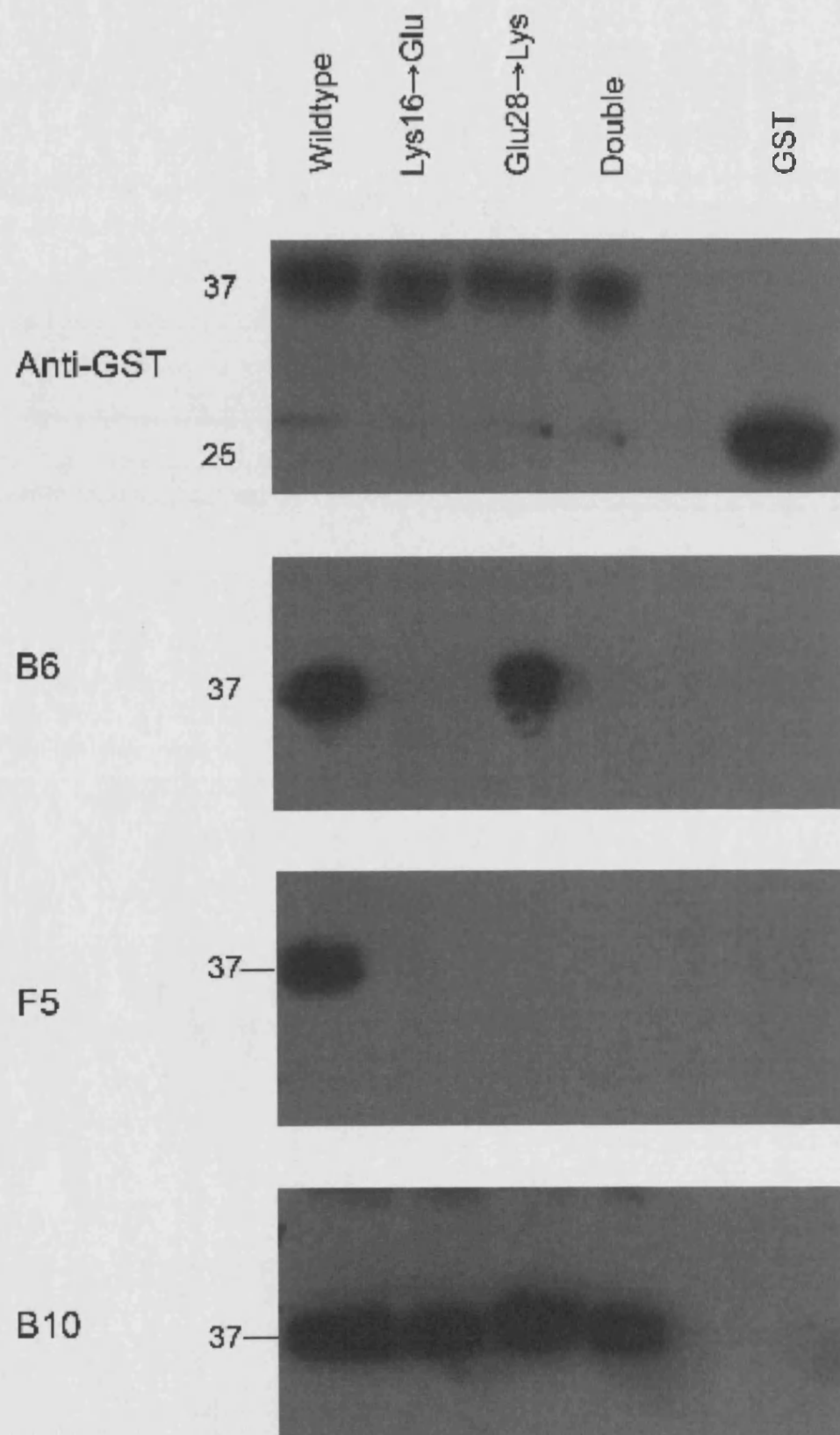


Figure 4.3: ELISA of B6 antibody binding to residues 16 and 28 single and double GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype, Lys16→Glu, Glu28→Lys, Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B6 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in orange, wildtype GST-MSP1₁₉ is shown in blue, Lys16→Glu GST-MSP1₁₉ variant is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in green and Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant is shown in pink.

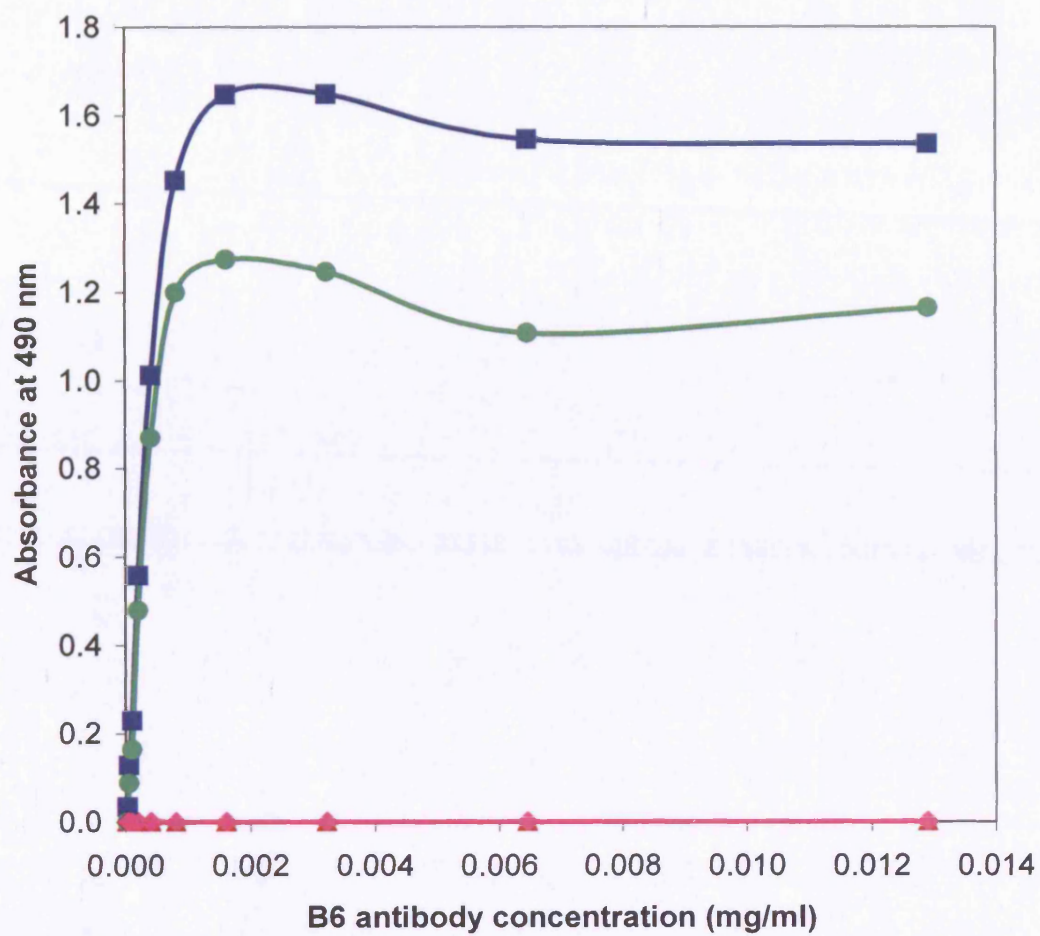


Figure 4.4: ELISA of F5 antibody binding to residues 16 and 28 single and double GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype, Lys16→Glu, Glu28→Lys, Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of F5 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in orange, wildtype GST-MSP1₁₉ is shown in blue, Lys16→Glu GST-MSP1₁₉ variant is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in green and Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant is shown in pink.

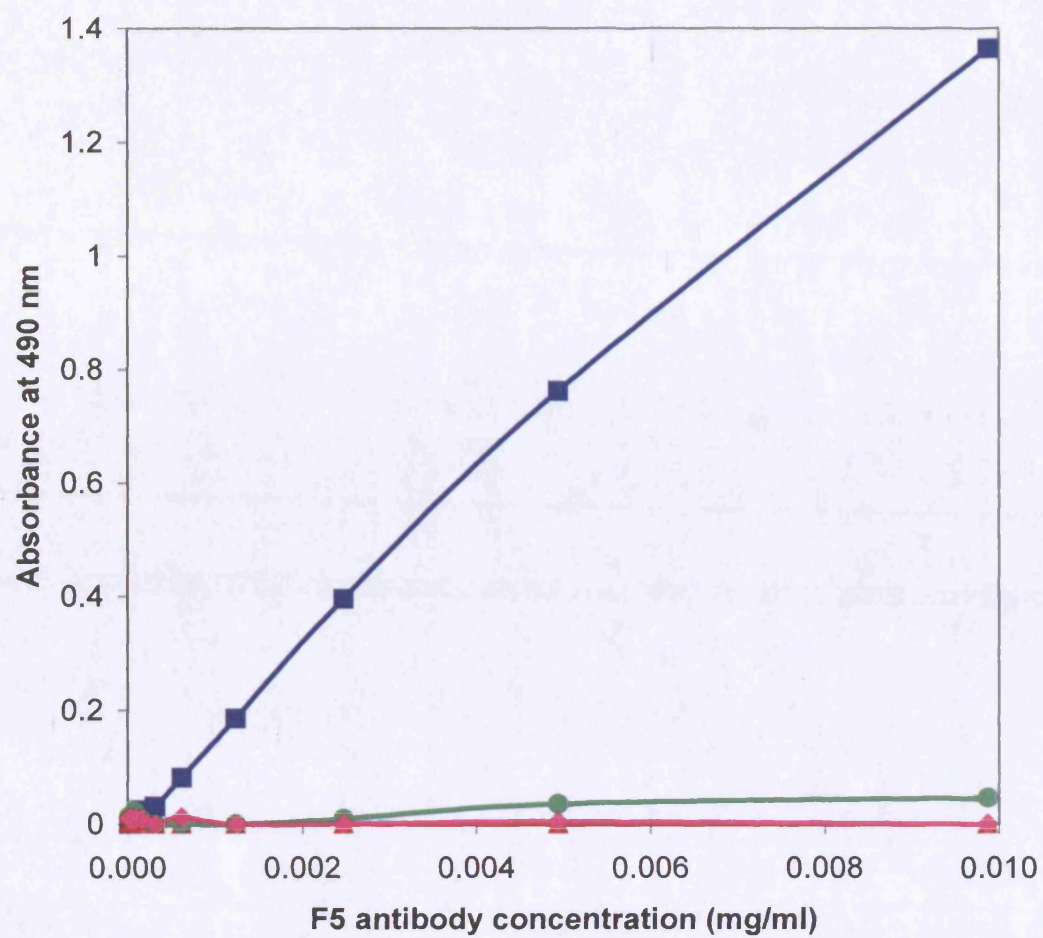
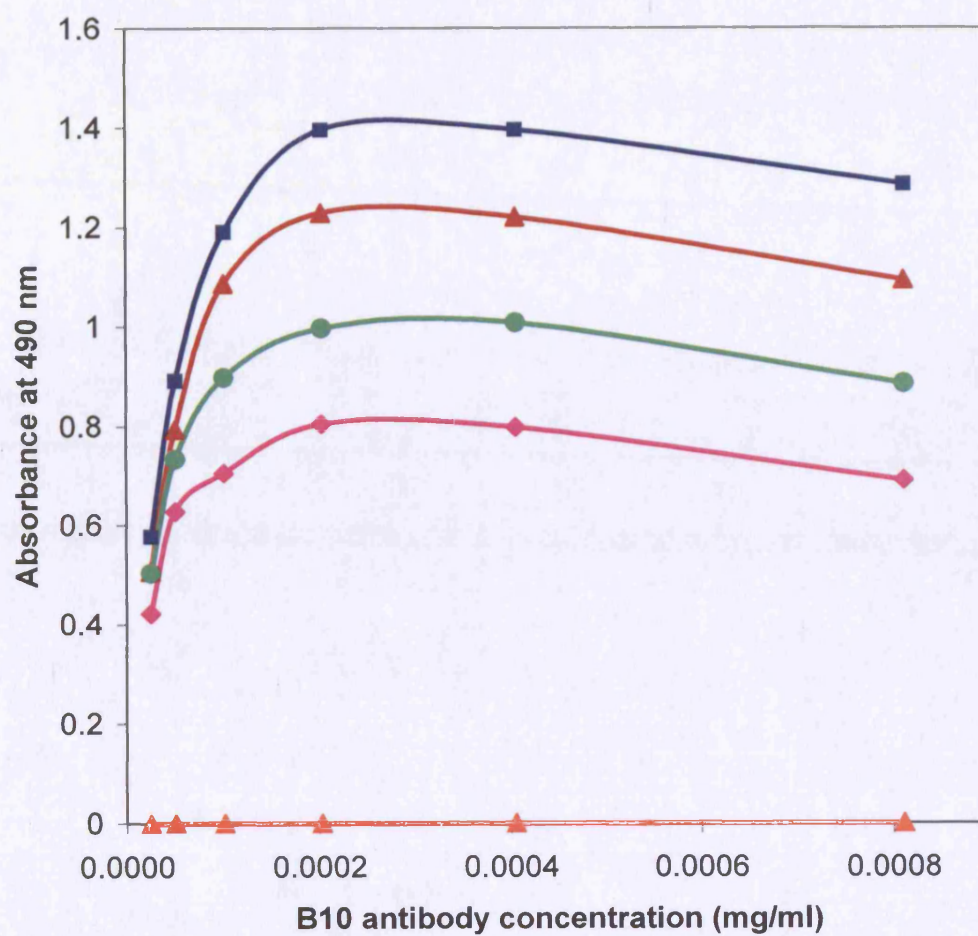


Figure 4.5: ELISA of B10 antibody binding to residues 16 and 28 single and double GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype, Lys16→Glu, Glu28→Lys, Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B10 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST-Tag less PBS control are shown on the graph. GST is shown in orange, wildtype GST-MSP1₁₉ is shown in blue, Lys16→Glu GST-MSP1₁₉ variant is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in green and Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant is shown in pink.



4.5 Discussion

The double Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant described in this chapter was designed to affect binding to B6, F5 and B10 monoclonal antibodies. The hypothesis for antibody binding to the double Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant was that it would have the combined effect of the single residues 16 and 28 variants. This would mean that it would not bind to B6 and F5 and would show reduced binding to B10. The western blotting and ELISA experiments presented here have proved the hypothesis correct as it has shown that the double Lys16→Glu/Glu28→Lys GST-MSP1₁₉ variant abolishes binding to B6 and F5 antibody and reduces B10 antibody binding. The reduction in B10 antibody binding shown in the ELISA experiments indicates an additive effect on the antibody binding with the double variant resulting in a larger reduction in the steepness of the antibody binding curve than the individual variants. The ELISA data suggests that both residues 16 and 28 are involved in B10 antibody binding and that by altering two residues that are involved in antibody binding it makes it more difficult for B10 to bind therefore resulting in a larger reduction in binding.

Chapter 5: Immunisation studies 1 – do the MSP1₁₉ variations affect protection?

5.1 Introduction

The amino acid variations in the MSP1₁₉ variant proteins discussed in chapters 3 and 4 showed an affect on binding to monoclonal antibodies *in vitro*. The three antibodies that were used in the binding studies in chapters 3 and 4 were produced by Spencer *et al.* (Spencer Valero *et al.*, 1998). Spencer *et al.* (Spencer Valero *et al.*, 1998) tested the ability of the monoclonal antibodies to suppress parasitaemia in mice during challenge infection with *P. yoelii* YM by passive immunisation with the monoclonal antibodies. These studies showed that B6 and F5 antibodies suppressed the challenge infection and that B10 antibody partially suppressed the challenge infection. These data could suggest that the MSP1₁₉ variants proteins may produce different antibodies *in vivo* to the wildtype protein.

In this chapter, I will discuss the immunisation studies that were designed to look at the affect of the amino acid variations on the ability of MSP1₁₉ to protect against challenge infection with the lethal *P. yoelii* YM parasite. Previous studies by Daly *et al.* (Daly & Long, 1993) showed that mice immunised with GST- MSP1₁₉ fusion proteins produced high titres of anti-MSP1 antibodies and the mice were protected from challenge infection with the lethal *P. yoelii* YM parasite. Ling *et al.* (Ling *et al.*, 1994) confirmed that MSP1₁₉ alone and GST-MSP1₁₉ was able to protect against challenge infection and that the conformation of the protein was important for protection. The immunisation studies discussed in this chapter are based on the experimental design of Ling *et al.* (Ling *et al.*, 1994).

The following hypothesis will be tested in the immunisation studies: there is one area that is important for binding of antibodies that protect against parasite challenge. There could be a number of outcomes to the immunisation studies: no difference in protection; a small difference in protection or no protection with the MSP1₁₉ variant proteins. If there is no difference in the protection between the wildtype and MSP1₁₉ variant

proteins, this could suggest that the individual amino acid changes are too small to have a significant effect on the immune response to the protein and the antibodies produced. This could also mean that the hypothesis is incorrect and that there are multiple sites that are important in the binding of antibodies that protect against parasite challenge. If there is a small difference in protection, this could suggest that the antibody response that is important in protection against parasite challenge is targeted at the region where the amino acid variation is and that the antibodies produced cannot recognise the native protein and therefore give protection. It could suggest that the amino acid change was too small to completely block protection and that some antibodies with the ability to protect against parasite challenge that are still able to recognise the native protein have been produced. It could again suggest that there could be multiple sites that are important in producing the protective immune response and that the change has affected one of them but antibodies produced to the other sites can still lead to protection. If there is no protection with the MSP1₁₉ variant proteins this could suggest that there is only one area that is important for the production of antibodies that protect against challenge infection and that the amino acid change to that area has altered the area enough to result in antibodies made to this area not recognising the native protein and not providing protection.

ELISA experiments will also be discussed in this chapter to compare antibody titres between the mice immunised with the wildtype and GST-MSP1₁₉ variant proteins because previous immunisation studies using MSP1₁₉ have suggested an important role of antibodies in the protection given by MSP1₁₉. The immunisation studies with MSP1₁₉ by Ling *et al.* (Ling *et al.*, 1994) showed the level of antibody produced against the parasite was highest in the mice that were protected from challenge infection (Ling *et al.*, 1994). It could therefore be hypothesized that the level of antibody against wildtype MSP1₁₉ may be lower for the mice immunised with the MSP1₁₉ variant proteins if they reduce the level of protection in the immunisation studies.

5.2 Immunisation studies with MSP1₁₉ variants

The immunisation studies were carried out as described in materials and methods section 2.3. Immunisation studies were carried out with the four single amino acid variants (Arg12→Leu, Lys16→Glu, Asn17→His, Glu28→Lys) and one double amino acid GST-MSP1₁₉ variant (Lys16→Glu/Glu28→Lys). Wildtype GST-MSP1₁₉ was used as a positive control and purified GST was used as a negative control. Following immunisation with the GST-MSP1₁₉ variants, the mice were challenged with *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily on Giemsa stained blood films. The percentage parasitaemia was determined to compare between the variants. The graphs in figures 5.1 to 5.7 show the parasitaemia for each mouse in the groups. Figure 5.1 shows the parasitaemia for the mice immunised with wildtype GST-MSP1₁₉. This shows that five of the mice immunised with wildtype GST-MSP1₁₉ were able to clear the parasites. Four of the mice had very low parasitaemia and one had higher parasitaemia. One of the mice was unable to clear the parasites and was killed by a schedule one method on day 7. Figure 5.2 shows the parasitaemia for the mice immunised with GST as negative controls. This shows a rapid increase in parasitaemia up to day 6 and all the mice were killed by a schedule one method on day 7. Figure 5.3 shows the parasitaemia for the mice immunised with Arg12→Leu MSP1₁₉ variant. This shows that four of the mice were able to clear the parasites and had very low parasitaemia. Two mice were unable to clear the parasites with one mouse starting with a low parasitaemia until day 10 and then a rapid increase. Figure 5.4 shows the parasitaemia for the mice immunised with Asn17→His MSP1₁₉ variant. This shows that three of the mice were able to clear the parasites. The other three mice did not clear the parasites and were killed by a schedule one method on day 7 although the parasitaemia of two of the mice were low (13 % and 39.5 %). Figure 5.5 shows the parasitaemia for the mice immunised with Lys16→Glu MSP1₁₉ variant. This shows that five of the mice were able to clear the parasites, three of the mice had low parasitaemia and two had higher parasitaemia with peak parasitaemia of 64.5 % and 46 %. One mouse was killed by a schedule one method on day 8 but had low parasitaemia of 19 %. Figure 5.6 shows the parasitaemia for the mice immunised with Glu28→Lys MSP1₁₉ variant. This shows that all of the mice had a rapid increase in parasitaemia and none of the mice were able

to clear the parasites. All of the mice were killed by a schedule one method, four on day 8 and two on day 12. Figure 5.7 shows the parasitaemia for the mice immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant. This shows that three mice had a rapid increase in parasitaemia and two were killed by a schedule one method by day 7 and the other on day 13. The other three mice had very high parasitaemia throughout the experiment with a peak at day 12. Figure 5.8 shows the average parasitaemia of the six mice in the groups. It shows that the overall patterns of parasitaemia for the wildtype, Arg12→Leu, Lys16→Glu and Asn17→His MSP1₁₉ variants were very similar with low parasitaemia. The overall patterns of parasitaemia for the Glu28→Lys and Lys16→Glu/Glu28→Lys MSP1₁₉ variants were different from the wildtype with larger increases in parasitaemia and higher parasitaemia.

5.3 ELISA analysis of antibody titres following immunisation with MSP1₁₉ variants

The antibody titres following immunisation with the GST-MSP1₁₉ variants were compared to those of the wildtype GST-MSP1₁₉ immunisations to determine whether any differences in protection from parasite challenge could be explained by differences in the level of antibody response to the GST-MSP1₁₉ variants. The ELISA experiments were carried out as described in materials and methods (section 2.3.1). In the ELISA experiments his-tagged wildtype MSP1₁₉ was used to analyse the antibody levels instead of GST-MSP1₁₉. This was to avoid problems associated with the production of antibodies to the GST portion of GST-MSP1₁₉ which could saturate the ELISA signal making it difficult to see small differences in antibody titre to the MSP1₁₉ portion. 1 µg/ml wildtype his-MSP1₁₉ was bound to the ELISA plate (the production of his-MSP1₁₉ is described in section 2.4). The proteins were probed with doubling dilutions of pooled serum samples from the six mice in the groups and 1/2000 dilution anti-mouse IgG HRP conjugate. The peroxidase was detected and absorbance was read at 490 nm. The ELISA results are shown in figure 5.9. The ELISA results show that all the mice immunised with all of the GST- MSP1₁₉ variants had produced antibodies to MSP1₁₉. The binding curve for the mice immunised with GST shows that the ELISA system used is specific for antibody binding to MSP1₁₉. The ELISA results show that

there is no significant difference between the antibody titres produced by the mice immunised with wildtype GST-MSP1₁₉ and the variant proteins. The results show a possible small reduction in the antibody titre for mice immunised with Glu28→Lys and double Lys16→Glu/Glu→Lys MSP1₁₉ variant.

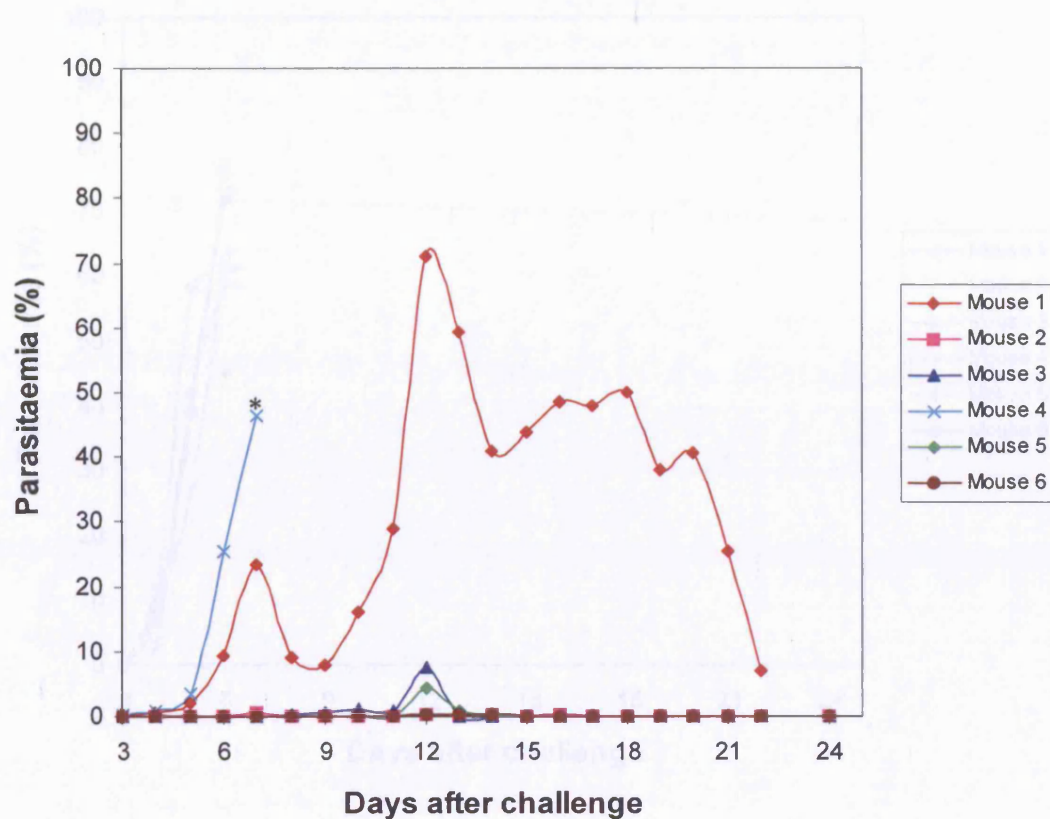


Figure 5.1: Course of *P. yoelii* YM infection in mice immunised with wildtype MSP1₁₉.

Six BALB/c mice were immunised with 10 µg of wildtype MSP1₁₉ in FCA followed by two injections with 40 µg of wildtype MSP1₁₉ in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

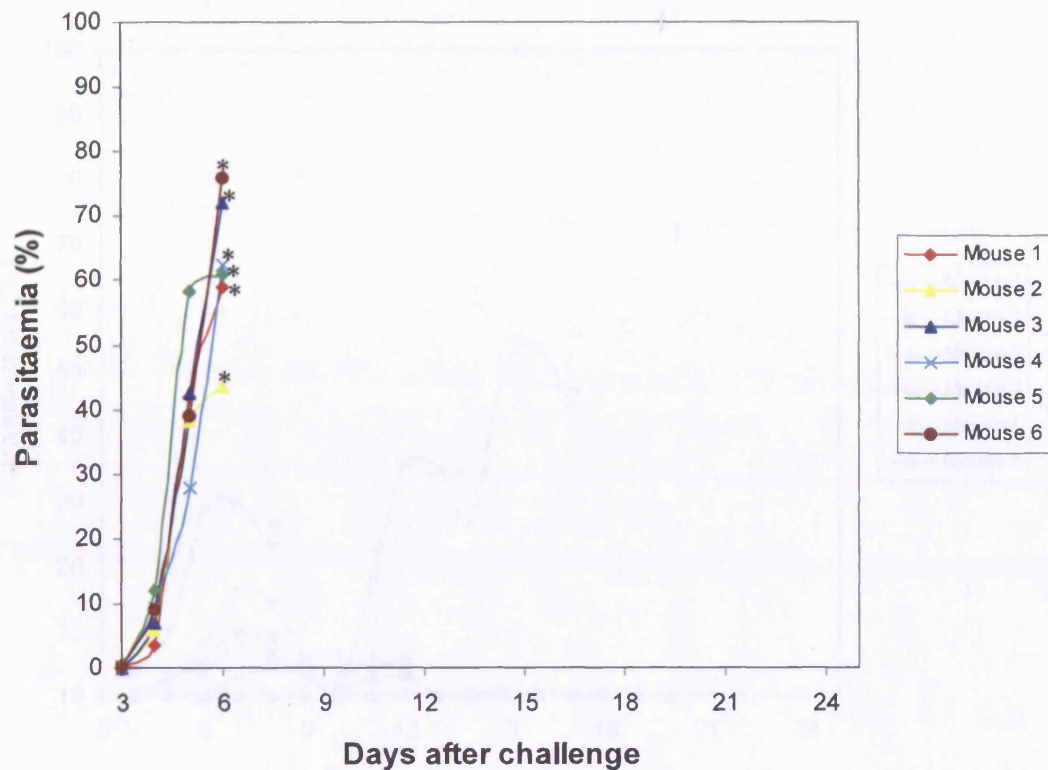


Figure 5.2: Course of *P. yoelii* YM infection in mice immunised with purified GST.

Six BALB/c mice were immunised with 10 µg of GST in FCA followed by two injections with 40 µg of GST in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

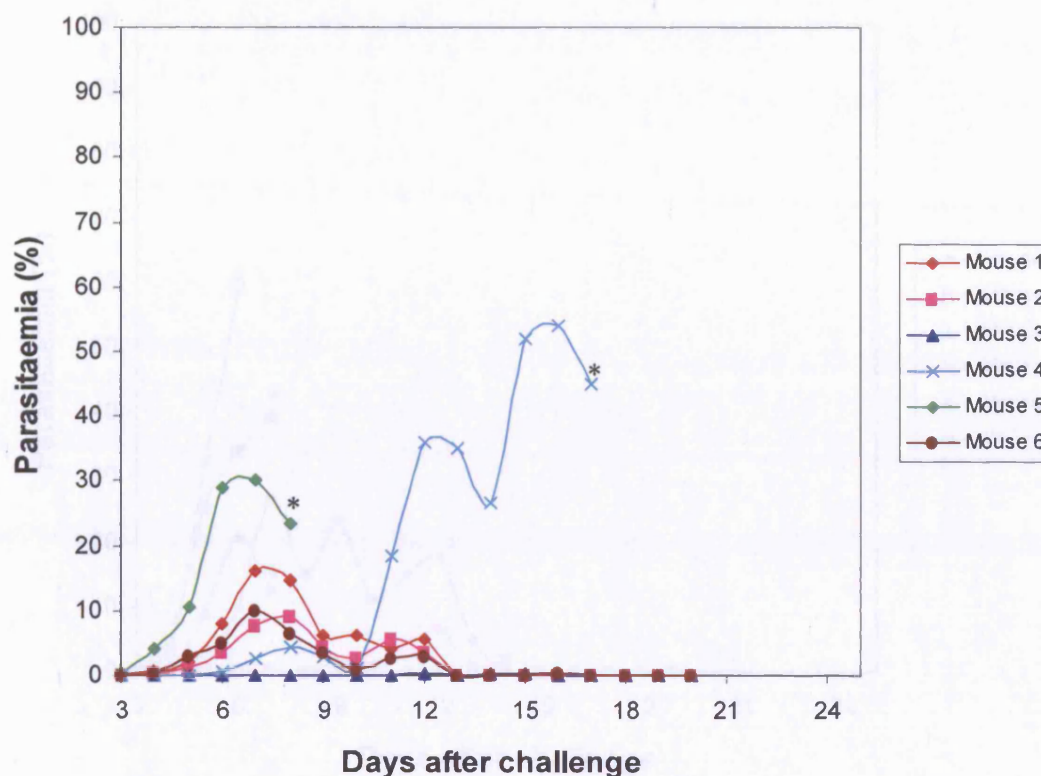


Figure 5.3: Course of *P. yoelii* YM infection in mice immunised with Arg 12→Leu MSP1₁₉ variant.

Six BALB/c mice were immunised with 10 µg of Arg12→Leu MSP1₁₉ variant in FCA followed by two injections with 40 µg of Arg12→Leu MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

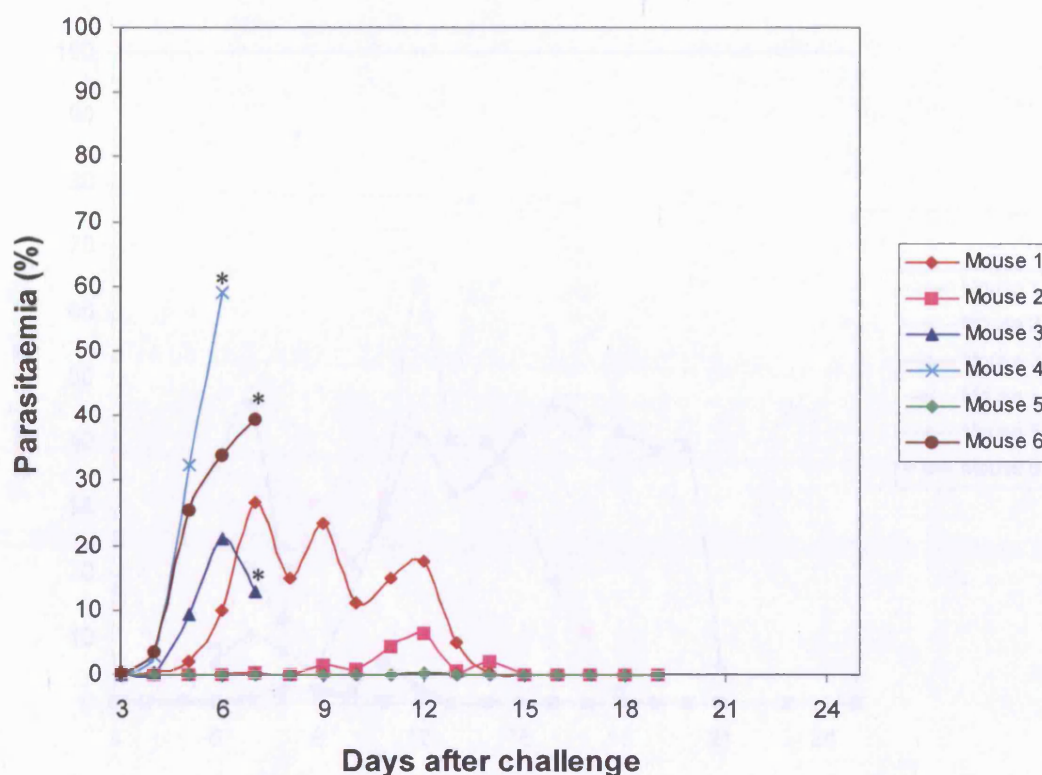


Figure 5.4: Course of *P. yoelii* YM infection in mice immunised with Asn17→His MSP1₁₉ variant.

Six BALB/c mice were immunised with 10 µg of Asn17→His MSP1₁₉ variant in FCA followed by two injections with 40 µg of Asn17→His MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

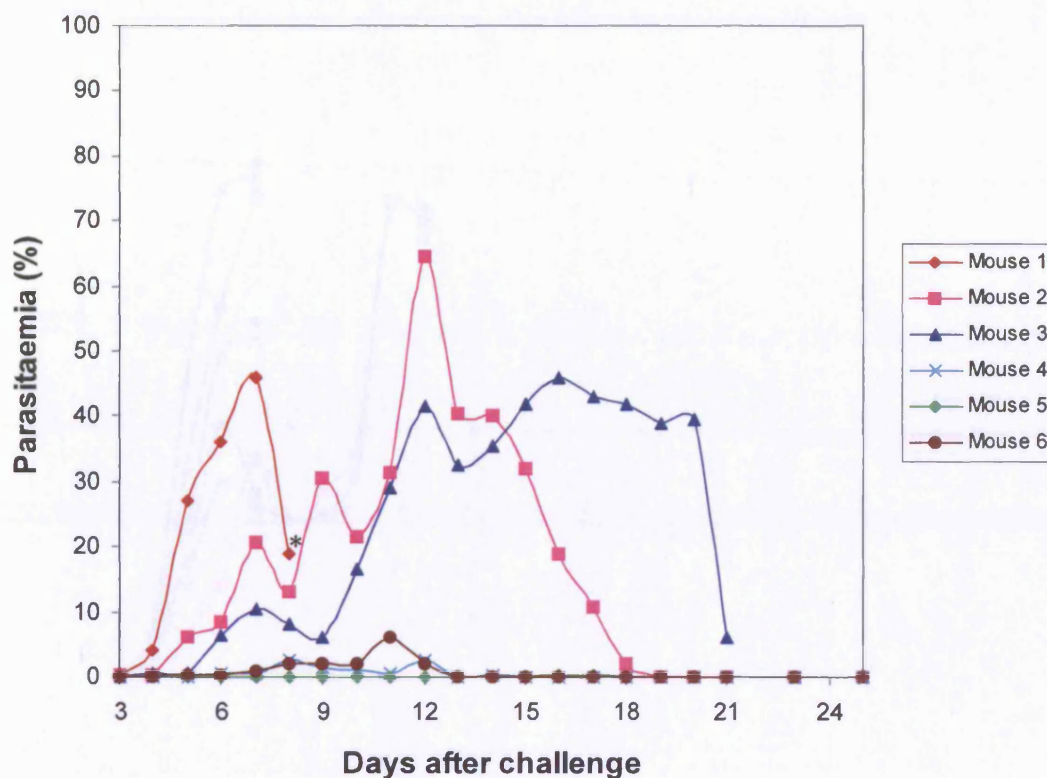


Figure 5.5: Course of *P. yoelii* YM infection in mice immunised with Lys16→Glu MSP1₁₉ variant.

Six BALB/c mice were immunised with 10 µg of Lys16→Glu MSP1₁₉ variant in FCA followed by two injections with 40 µg of Lys16→Glu MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

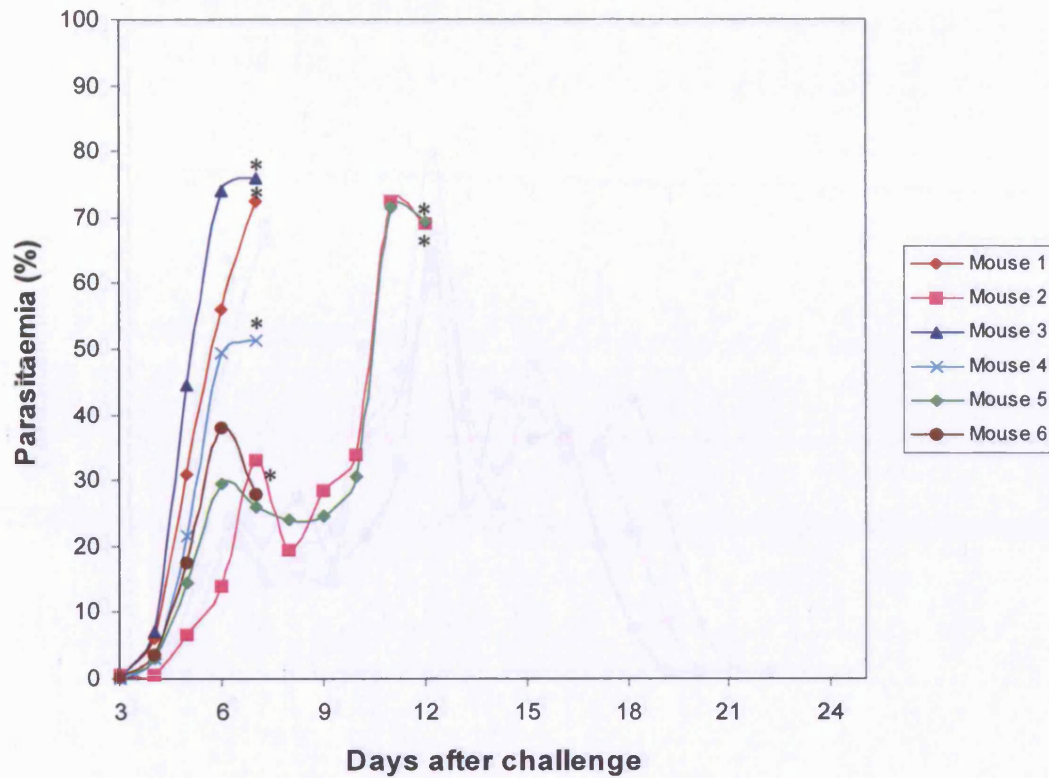


Figure 5.6: Course of *P. yoelii* YM infection in mice immunised with Glu28→Lys MSP1₁₉ variant.

Six BALB/c mice were immunised with 10 µg of Glu28→Lys MSP1₁₉ variant in FCA followed by two injections with 40 µg of Glu28→Lys MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

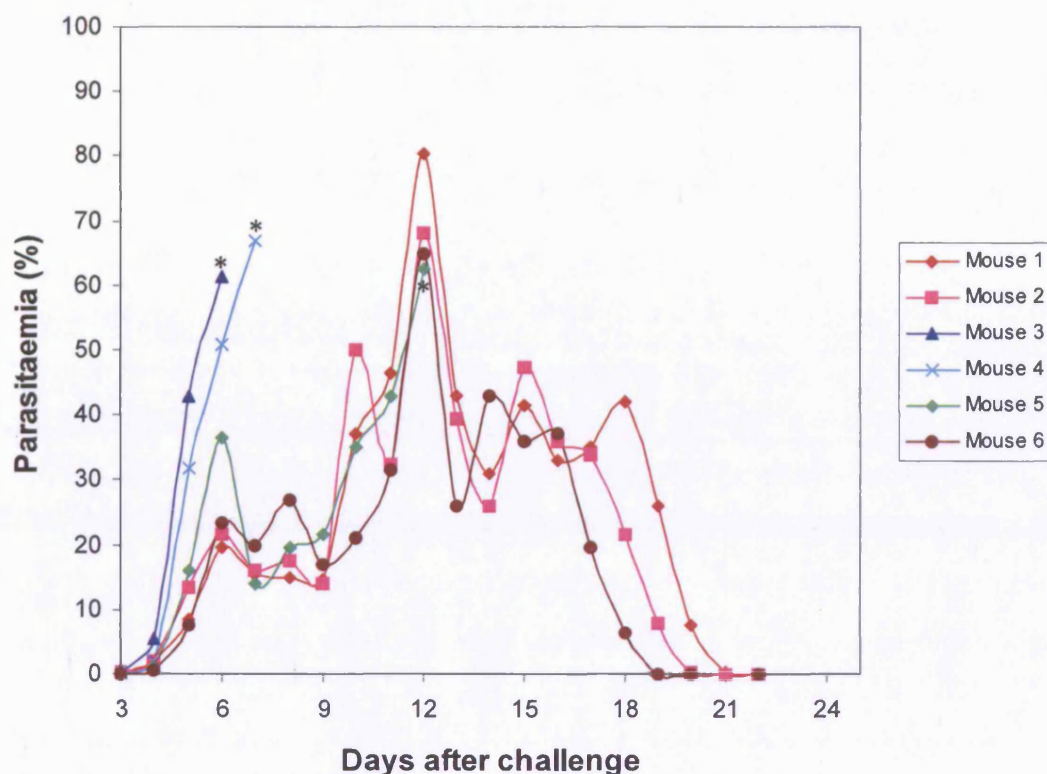


Figure 5.7: Course of *P. yoelii* YM infection in mice immunised with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant.

Six BALB/c mice were immunised with 10 µg of double Lys16→Glu/Glu28→Lys MSP1₁₉ variant in FCA followed by two injections with 40 µg of double Lys16→Glu/Glu28→Lys MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

Figure 5.8: Course of *P. yoelii* YM infection in groups of mice immunised with wildtype and MSP1₁₉ variants.

Groups of six BALB/c mice were immunised with 10 µg of wildtype or MSP1₁₉ variants or GST in FCA followed by two injections with 40 µg of protein in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The average percentage parasitaemia for the groups are plotted on the graph. The average parasitaemia for mice immunised with wildtype MSP1₁₉ is shown in red, with GST is shown in pink, with Arg12→Leu MSP1₁₉ variant is shown in black, with Asn17→His MSP1₁₉ variant is shown in purple, with Lys16→Glu MSP1₁₉ variant is shown in cyan, with Glu28→Lys MSP1₁₉ variant is shown in blue and with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in orange. Asterisks indicate when a mouse died or was killed by a schedule one method.

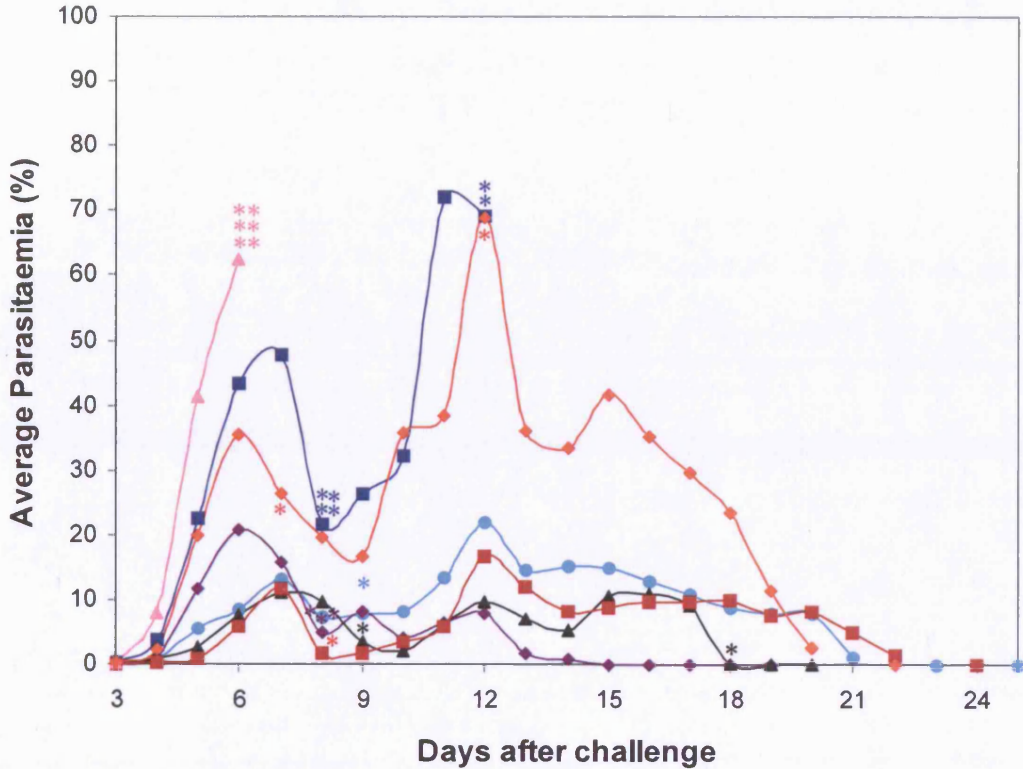
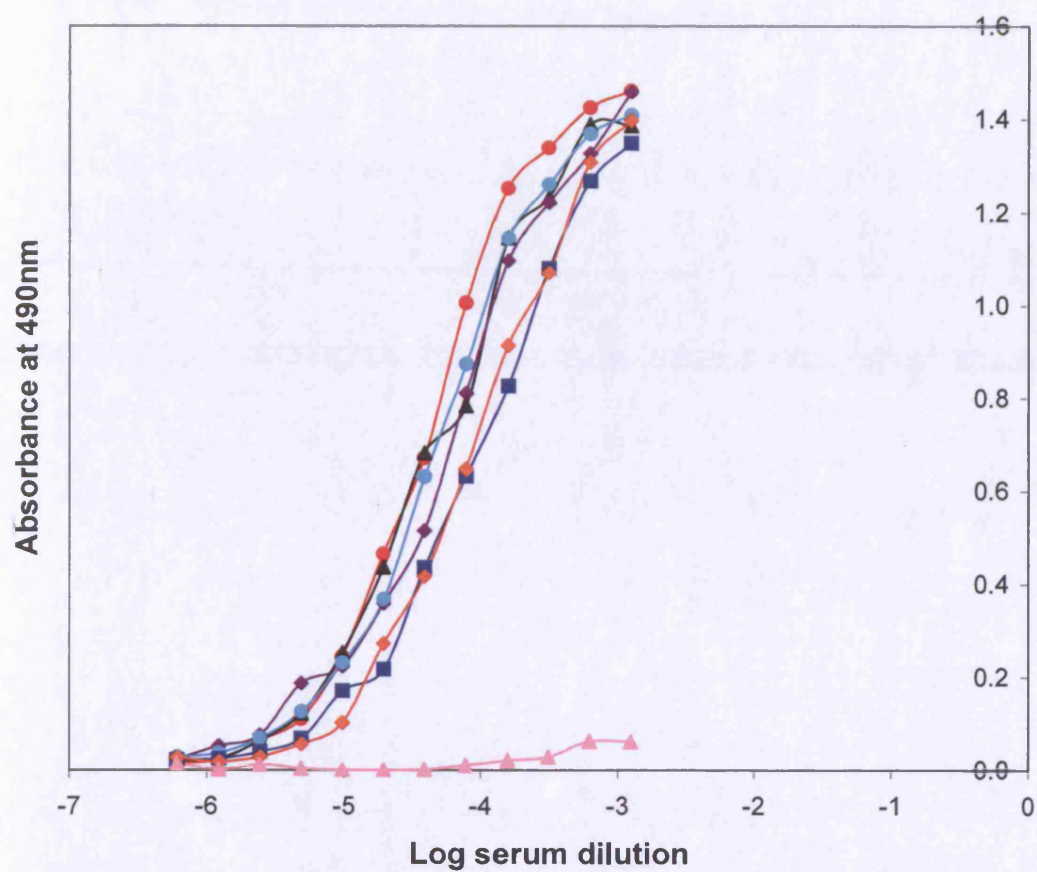


Figure 5.9: Antibody binding curves for pooled serum from groups of mice immunised with wildtype and MSP1₁₉ variants against his-tagged wildtype MSP1₁₉.

1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of serum samples from the mice immunised with wildtype and MSP1₁₉ variants and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. The serum samples for the six mice in the groups were pooled together. Serum from mice immunised with GST is shown in pink, with wildtype MSP1₁₉ is shown in red, with Lys16→Glu MSP1₁₉ is shown in cyan, Asn17→His MSP1₁₉ is shown in purple, Arg12→Leu MSP1₁₉ is shown in black, Glu28→Lys MSP1₁₉ is shown in blue and Lys16→Glu/Glu28→Lys is shown in orange.



5.4 Discussion

The immunisation studies presented in this chapter, have suggested that there are differences between the protection observed following immunisation with the wildtype protein and with the variants. The overall results for the mice immunised with GST has shown that the GST-portion of the GST-MSP1₁₉ is not protecting the mice from challenge infection. This confirms that the protection seen in the experiment for the other groups is a result of the MSP1₁₉ portion. The overall results have shown that the immunisation with wildtype protein protects against parasite challenge because the parasitaemia is low and the mice clear the parasite. The protection and low parasitaemia seen in these immunisation studies for the wildtype GST-MSP1₁₉ is in agreement with the protection observed by Daly *et al.* (Daly & Long, 1993) and Ling *et al.* (Ling *et al.*, 1994). Immunisation with Arg12→Leu, Lys16→Glu and Asn17→His MSP1₁₉ variants protects against parasite challenge and mice clear the parasite in a similar way to immunisation with the wildtype protein. Immunisation with Glu28→Lys and double Lys16→Glu/Glu28→Lys has a significant affect on protection compared to the protection seen for the wildtype protein. Immunisation with the Glu28→Lys MSP1₁₉ variant does not protect against parasite challenge this suggests that residue 28 is in the important area of the protein for the production of antibodies that protect against parasite challenge. This could mean that when the mice are immunised with Glu28→Lys MSP1₁₉ variant they either do not produce antibodies to the area of the protein containing residue 28 or they produce antibodies to this area but the difference to residue 28 means that the antibodies are not able to bind to the wildtype MSP1₁₉. This would mean that on challenge infection the antibodies produced would be unable to provide a protective immune response. The results for the Glu28→Lys MSP1₁₉ variant could agree with the hypothesis for the immunisation studies that there is one area that is important for the binding of antibodies that protect against challenge infection and by changing residue 28 in this area the response produced no longer protects against parasite challenge.

The parasitaemia counts for the individual mice showed that there was some variation between the individual mice in the group. This may have been because of variations in the experiment that could not be controlled. The mice used in the immunisation study

were cousins because there were too many mice required for just brothers and sisters to be used. This could have introduced some genetic variation between the mice which could have resulted in differences in immunity. Small differences in the amount of antigen the mice were immunised with could have occurred and this could have influenced the level of immune response. Differences in the uptake of the antigen between the individual mice following immunisation could have influenced the immune response produced. The behaviour of the individual mice may have influenced the progression of malaria, for example if a mouse was not eating or drinking as much as the other mice it could become unwell more quickly and be less able to fight the disease. The method of measuring the parasitaemia could also have introduced inaccuracies in the experiment because it relied on one person visually counting slides. If the mice were anaemic this could have lead to inaccuracies in the counts because the blood smears from the anaemic mice did not give an even coverage of blood cells and cells appeared in clumps. This could have lead to the blood cells that were counted to not be representative of the overall parasitaemia. The parasites could have sequestered in the spleen, brain or other organs and this would not be taken into account by counting the parasites in the peripheral blood and could result in lower counts and inaccuracies if variant proteins had altered the level of sequestration. Additional errors could have been introduced in the decisions regarding when to kill the mice by a schedule one method as some mice became unwell at very low parasitaemia levels while other mice were well at high parasitaemia. For example, the results for the individual mice immunised with Asn17→His MSP1₁₉ variant shows that 3 mice cleared the parasites and the other 3 mice were killed by schedule one method on day 7 with relatively low parasitaemia levels.

The ELISA results show that there is no significant difference between the antibody titres to wildtype MSP1₁₉ produced by the wildtype and MSP1₁₉ variant proteins. There is a small possible reduction for the Glu28→Lys MSP1₁₉ and the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant proteins. This suggests that the overall antibody titre may not be that important for protection against challenge infection. This disagrees with the immunisation studies of Ling *et al.* (Ling *et al.*, 1994) where the antibody response to the parasite was highest in the mice that were protected against parasite challenge. The ELISA data presented in this chapter suggested that the overall

antibody titre may not be that important in determining if the protein protects against challenge infection but could suggest that instead the fine specificity of the antibody produced may be the most important factor in determining protection.

Chapter 6: Structural analysis of MSP1₁₉ variants

6.1 Introduction

In chapters 3-5 the affects of the amino acid variations to MSP1₁₉ on the ability of it to bind to antibodies has been investigated. The data has suggested that the amino acid variations at residues 12, 16, 17 and 28 are important for antibody binding *in vitro* and for residue 28 *in vivo*. Any changes to the amino acid composition and sequence of a protein, even changing just a single amino acid can potentially affect the wider 3D structure of the protein. It is therefore important to confirm whether the individual amino acid changes made in this project have affected the overall 3D structure of the protein.

In this chapter, I am going to describe the approach taken to identify structural differences between the wildtype MSP1₁₉ and the MSP1₁₉ variants. There was no structure available for *P. yoelii* MSP1₁₉ but there were structures available for MSP1₁₉ from other species. The first approach that was taken to analyse the structural differences was to create a homology model of *P. yoelii* MSP1₁₉ using the structures that were available in the Protein Databank for MSP1₁₉ structures from other species. The second approach that was taken to analyse structural differences in more detail was to carry out ¹⁵N-HSQC NMR analysis. In order to carry out the NMR analysis his-tagged proteins were created using *P. pastoris* to obtain milligram quantities of pure protein.

6.2 Molecular modelling of *P. yoelii* MSP1₁₉

The homology model of *P. yoelii* MSP1₁₉ was created using the Swiss Model Homology Modelling Server (Guex & Peitsch, 1997, Schwede *et al.*, 2003). The Swiss Model Homology Modelling Server uses an automated approach to create a model of a protein based on structures of homologous proteins that are available in the protein databank.

The modelling procedure consists of four steps (Schwede *et al.*, 2003) as described in table 6.1.

Table 6.1: The steps carried out by the Swiss Model Homology Modelling Server to create a homology model.	
Step	Procedure
1	Selection of template sequences from PDB.
2	Alignment of up to 5 template sequences against the target sequence.
3	Building the model based on average backbone atom positions of the templates and then building the side chains.
4	Evaluation of the model and energy minimisation.

The homology model is shown in the top panel of figure 6.1. The positions of the amino acid variations were identified using Deepview/Swiss Pdb viewer and RasTop. The model suggested residues 12, 16, 17 and 28 were found on the surface of the protein. In order to predict the effects of changing the amino acids on the protein, the amino acids were changed *in silico* using Deepview/Swiss Pdb viewer. The *in silico* variations are shown in figure 6.1. This figures shows that there is a small difference in the size of the residues in the Arg12→Leu and Glu28→Lys variations. The Lys16→Glu variation gives a residue that sticks out much less in the optimum rotamer conformation and the Asn17→His variation gives a larger residue that sticks out more. For the Glu28→Lys variation there appears to be a difference in the angle of the side chain when looking at the 270° rotation. When Glutamic acid 28 is changed to lysine in the model, the lysine 28 residue appears to be angled downwards away from residue 12 whereas in the wildtype model it appears to be below residue 12.

6.3 Expression and purification of his-MSP1₁₉ variants

In order to carry out NMR analysis at least 5 mg of purified protein was required for each variant. The expression of GST-MSP1₁₉ in *E. coli* (described in chapter 3) resulted in microgram quantities of protein. The GST-tag on the protein is around 25 kDa in size and contains many different amino acid residues. The GST-tagged MSP1₁₉ variant proteins were therefore unsuitable for use in NMR analysis because the large tag would create lots of NMR signals that would be hard to distinguish from those belonging to the protein. The overall size of the GST-tagged protein would also cause problems in obtaining good NMR spectra.

In order to make milligram quantities of protein, the MSP1₁₉ variants were expressed in *P. pastoris* as his-tagged proteins. The his-tag was added to the gene construct to allow for easy purification of the proteins using nickel resin. The his-tag would not interfere with the NMR analysis as it only contains 6 histidine residues that could be identified in the NMR spectra as they would form peaks in the same area as histidines that were not in folded proteins.

The *P. pastoris* system was chosen because it has been shown in the literature to give high yields of proteins (Romanos *et al.*, 1992, Sreekrishna *et al.*, 1997) including human EGF and murine EGF (Clare *et al.*, 1991). The proteins are expressed from the alcohol oxidase gene (AOX1) which catalyses the first step in the metabolism of methanol. When methanol is added to the cells, the AOX1 promoter induces protein production to give high levels of expression of recombinant protein. The AOX1 gene has a very strong promoter which means that when the cells are grown in methanol up to 30 % of total soluble proteins are AOX1 protein products. The pPIC9K plasmid was chosen for expression in *Pichia pastoris* because this vector gives expression of secreted proteins in the medium via the α -factor signal sequence. *P. pastoris* secretes very low levels of native proteins into the culture medium which means that by expressing the MSP1₁₉ proteins as secreted proteins it will help make the purification of the protein easier by limiting the amount of protein contaminants present before purification.

The gene for MSP1₁₉ is very AT-rich. This makes it difficult to express in *P. pastoris*. Initial expression studies were carried out with the native wildtype MSP1₁₉ and MSP1₁₉ variant genes. This gave little or no protein expression (data not shown). In order to obtain high yields of protein in *P. pastoris* the genes for the MSP1₁₉ proteins were recodonised (as described in section 2.4.2) to increase the GC content as this has been shown by Woo *et al.* to increase protein yields (Outchkourov *et al.*, 2002, Woo *et al.*, 2002).

The recodonised MSP1₁₉ variants genes were cloned into pPIC9K (as described in section 2.4). Small scale (50 ml) expression tests were carried out with the MSP1₁₉ variants over 96 hours to determine the expression time that gives the best yield of protein with the lowest amount of protein breakdown and contaminants. The results for the small scale expression test are shown in figure 6.2. The MSP1₁₉ variants proteins appear as clear bands between 15 and 20 kDa. This shows that all of the MSP1₁₉ variants have been expressed at all the time points. It suggests that the best time for protein expression is 72 hours as the level of protein increases up to 72 hours and there is no significant increase after 72 hours at the 96 hour time point.

The MSP1₁₉ variant proteins from the small scale expression test were purified using Ni-NTA agarose. This is shown in figure 6.3 for the wildtype protein (panel A) and Glu28→Lys MSP1₁₉ variant (panel B). This shows that the protein has been purified successfully using the Ni-NTA (shown in lane 6). It shows a low molecular weight contaminant present after the purification. This contaminant was removed during the dialysis step. The protein would be pure enough for NMR analysis.

For NMR analysis to compare the wildtype and MSP1₁₉ variant proteins, ¹⁵N labelled protein was required. ¹⁵N labelled protein was produced by substituting the ammonium sulphate in the culture medium with ¹⁵N labelled ammonium sulphate. The ammonium sulphate in the culture medium was providing the sole nitrogen source for the production of the protein, by substituting it with ¹⁵N labelled ammonium sulphate the protein produced would incorporate ¹⁵N instead of ¹⁴N into the proteins. Large scale expression of the MSP1₁₉ variants was carried out to produce the 5 mg of pure protein required for NMR analysis. The large scale expression and purification of wildtype

(panel A) and Glu28→Lys MSP1₁₉ variant (panel B) is shown in figure 6.4. This confirms that the ¹⁵N labelled proteins were expressed and purified. The other MSP1₁₉ variants were expressed and purified in the same way as wildtype and Glu28→Lys MSP1₁₉ variant (data not shown).

6.4 ¹⁵N-HSQC NMR analysis of his-MSP1₁₉ variants

MSP1₁₉ is around 12 kDa. One dimensional ¹H-NMR spectra would therefore have a very large number of overlapping peaks for all of the protons in the protein. Figure 6.5 shows the 1D ¹H-NMR spectra for the wildtype (panel A) and Glu28→Lys variant (panel B). This shows a very large number of overlapping peaks for both structures and shows that the overall shape of the spectra is the same for both proteins. This can only give limited information about the proteins and does not allow the individual amino acids that have moved to be identified. Labelling the proteins with ¹⁵N can overcome this problem because it is possible to take advantage of the larger chemical shift dispersion than ¹H (Norwood *et al.*, 1989) and run 2D and 3D NMR experiments. The ¹⁵N labelled proteins were used in 2D ¹⁵N-HSQC spectroscopy experiments. HSQC stands for Heteronuclear Single Quantum Correlation spectroscopy. In the ¹⁵N-HSQC spectra, there is a peak for every ¹H bonded directly to ¹⁵N. This means that a peak will be seen for the backbone NH of every amino acid in the protein apart from proline. The side chain NH for arginine, asparagine and glutamine will also be seen. The position of the peak will depend upon the chemical environment of the amino acid residue. This will depend upon the type of amino acid residue, the conformation of the protein backbone and the nearby amino acid residues.

The position of the peaks can identify if a protein is folded. If the amino acids were in a random coil conformation, each type of amino acid would be chemically equivalent. This would mean that each type of residue would give a peak in a particular position on the spectra. If the amino acids are in a folded protein, the individual amino acids are chemically non-equivalent as their chemical shifts will depend on the conformation of the protein backbone and the nearby amino acid residues. The overall shape of the spectrum will therefore confirm if the MSP1₁₉ variants are still folded. If the proteins are folded the peaks will be spread out across the spectrum. If the protein is unfolded

the chemical shifts will be equivalent for each type of amino acid and only a few signals will be seen instead of lots of spread out signals.

The positions of the individual ¹⁵N-¹H peaks can be compared for the wildtype and the MSP1₁₉ variants. If a peak has moved, this will indicate that this amino acid is in a different chemical environment. This will suggest that this area of the molecule has changed. The main limitation of the ¹⁵N-HSQC experiment is that the distance moved in the ¹⁵N or ¹H dimension cannot be directly correlated to a movement in angstroms. It can only suggest the amount of chemical perturbation of that residue. A schematic representation of a ¹⁵N-HSQC spectrum is shown in figure 6.6 (panel A) with the assigned wildtype MSP1₁₉ ¹⁵N-HSQC spectra (panel B) (assignment of the wildtype MSP1₁₉ spectra will be discussed in chapter 9).

¹⁵N-HSQC spectra were acquired at 25 °C for each of the MSP1₁₉ variants as described in section 2.5.2 of materials and methods. The spectra were overlayed on top of the wildtype MSP1₁₉ spectra using SPARKY software (Goddard & Kneller). The spectra were compared to identify NH peaks that had moved. For the Arg12→Leu, Lys16→Glu and Asn17→His variants, the spectra were assigned based on the wildtype spectrum. For the peaks that had moved, it was assumed that the peak would correspond to the assignment of the closest wildtype peak. This approach was used because very few residues had moved which meant there were few ambiguities as to where the peaks had moved to. For the Glu28→Lys variant, the spectrum was assigned independently of the wildtype spectrum and 3D ¹³C and ¹⁵N NMR experiments (as discussed in chapter 9). The double Lys16→Glu/Glu28→Lys variant spectrum was assigned based on the wildtype and the Glu28→Lys variant spectra. The chemical shift differences in the ¹H and ¹⁵N dimension were combined using the formula:

$$\text{Absolute } ((\text{variant } ^{15}\text{N chemical shift} - \text{wildtype } ^{15}\text{N chemical shift}) \div 5) + \text{Absolute } (\text{variant } ^1\text{H chemical shift} - \text{wildtype } ^1\text{H chemical shift}) \div 2$$

The combined chemical shift differences were plotted on bar charts. Figure 6.7 shows the spectrum for Arg12→Leu variant in black overlayed on top of the wildtype spectrum shown in red (top panel). This shows that overall, the spectra were the same and that the very few NH peaks had moved. The peaks that had moved in the Arg12→Leu variant spectrum can be seen in the histogram (bottom panel). This

indicates that there was very little movement in this spectrum compared to the wildtype and indicates that the largest movement seen was for residue 13 NH and there were no large clusters of movement. Figure 6.8 shows the spectrum for Lys16→Glu variant in cyan overlayed on top of the wildtype spectrum shown in red (top panel). This shows that overall the chemical shifts of most of the residues are unchanged between the two spectra. The peaks that had moved in the Lys16→Glu variant spectrum can be seen in the histogram (bottom panel). This indicates that there was very little movement in this spectrum compared to the wildtype and indicates that the largest movement seen was for residue 16 NH itself and there were no large clusters of movement. Figure 6.9 shows the spectrum for Asn17→His variant in purple overlayed on top of the wildtype spectrum shown in red (top panel). Similarly, this shows that overall the spectra were the same and that very few NH peaks have moved as seen in the histogram (bottom panel). The largest movement seen is for residue 17 NH itself and there are no large clusters of movement. Figure 6.10 shows the spectrum for Glu28→Lys variant in blue overlayed on top of the wildtype spectrum shown in red (top panel). It is clear that the two spectra are quite different. In contrast to Arg12→Leu, Lys16→Glu and Asn17→His there were a large number of NH peaks that had moved in both the ¹⁵N and ¹H dimensions. The peaks that had moved in the Glu28→Lys variant spectrum can be seen in the histogram (bottom panel). This indicates that there are a number of clusters of movement including between residues 5 and 14, 19 and 22, 27 and 31 and at the end of the protein around residue 95. Figure 6.11 shows the spectrum for the double Lys16→Glu/Glu28→Lys variant in gold overlayed on top of the wildtype spectrum shown in red (top panel). This shows a number of NH peaks had moved in both the ¹⁵N and ¹H dimension. The peaks that had moved in the double Lys16→Glu/Glu28→Lys variant spectrum can be seen in the histogram (bottom panel). This indicates that there are a number of clusters of movement including between residues 5 and 14, 19 and 22, 27 and 31 and at the end of the protein around residue 95. Figure 6.12 shows the spectrum of the double Lys16→Glu/Glu28→Lys variant (shown in gold) overlayed on top of the Glu28→Lys variant spectrum (shown in blue). This indicates that these two spectra are very similar suggesting that the majority of the movement seen for the double Lys16→Glu/Glu28→Lys variant can be seen with just the Glu28→Lys variation. Figure 6.13 shows a histogram comparing the combined chemical shift difference for Lys16→Glu variant (shown in cyan), Glu28→Lys variant (shown in

blue) and double Lys16→Glu/Glu28→Lys variant shown in gold. This indicates that the majority of the chemical shift differences seen in the double Lys16→Glu/Glu28→Lys variant are from the addition of the chemical shift differences for the individual residue 16 and 28 variations. It also shows that there are no additional differences in the double Lys16→Glu/Glu28→Lys variant that do not appear in the individual variants.

Figure 6.14 shows the residues corresponding to the NH peaks movements (of 0.05 ppm or more) in the spectra highlighted on the 3D structure of wildtype *P. yoelii* MSP1₁₉ (the production of the 3D structure of wildtype *P. yoelii* MSP1₁₉ will be discussed in chapter 9). This shows that the NH peaks that have moved in the Arg12→Leu variant spectrum correspond to residues that are in direct contact with residue 12. For the Lys16→Glu variant there are no NH peaks apart from residue 16 that have moved more than 0.05 ppm. The NH peaks that have moved in the Asn17→His variant spectrum are located in the same half of the first EGF domain as residue 17 but are not all in direct contact with residue 17. The NH peaks that have moved in the Glu28→Lys variant spectrum correspond to residues that are covering a large part of the first EGF domain and part of the second EGF domain and include residues that are not in contact with residue 28. The NH peaks that have moved in the double Lys16→Glu/Glu28→Lys variant spectrum correspond to residues that are covering a large part of the first EGF domain and part of the second EGF domain in the same way as the Glu28→Lys variant. The NH peaks that have moved in the double Lys16→Glu/Glu28→Lys variant spectrum also correspond to residues close to residue 16.

Figure 6.1: *In silico* variation of residues in *P. yoelii* MSP1₁₉ model.

The amino acid residues that were altered experimentally were altered *in silico* on the *P. yoelii* MSP1₁₉ model using Deepview/Swiss Pdb viewer. The wildtype *P. yoelii* MSP1₁₉ model is shown in the top panel and the models with the amino acids altered are shown in the lower panels. Arg12→Leu is shown in green, Lys16→Glu is shown in pink, Asn17→His is shown in yellow and Glu28→Lys is shown in blue. The C-terminal residue is shown in black and the N-terminal residue is shown in light blue.

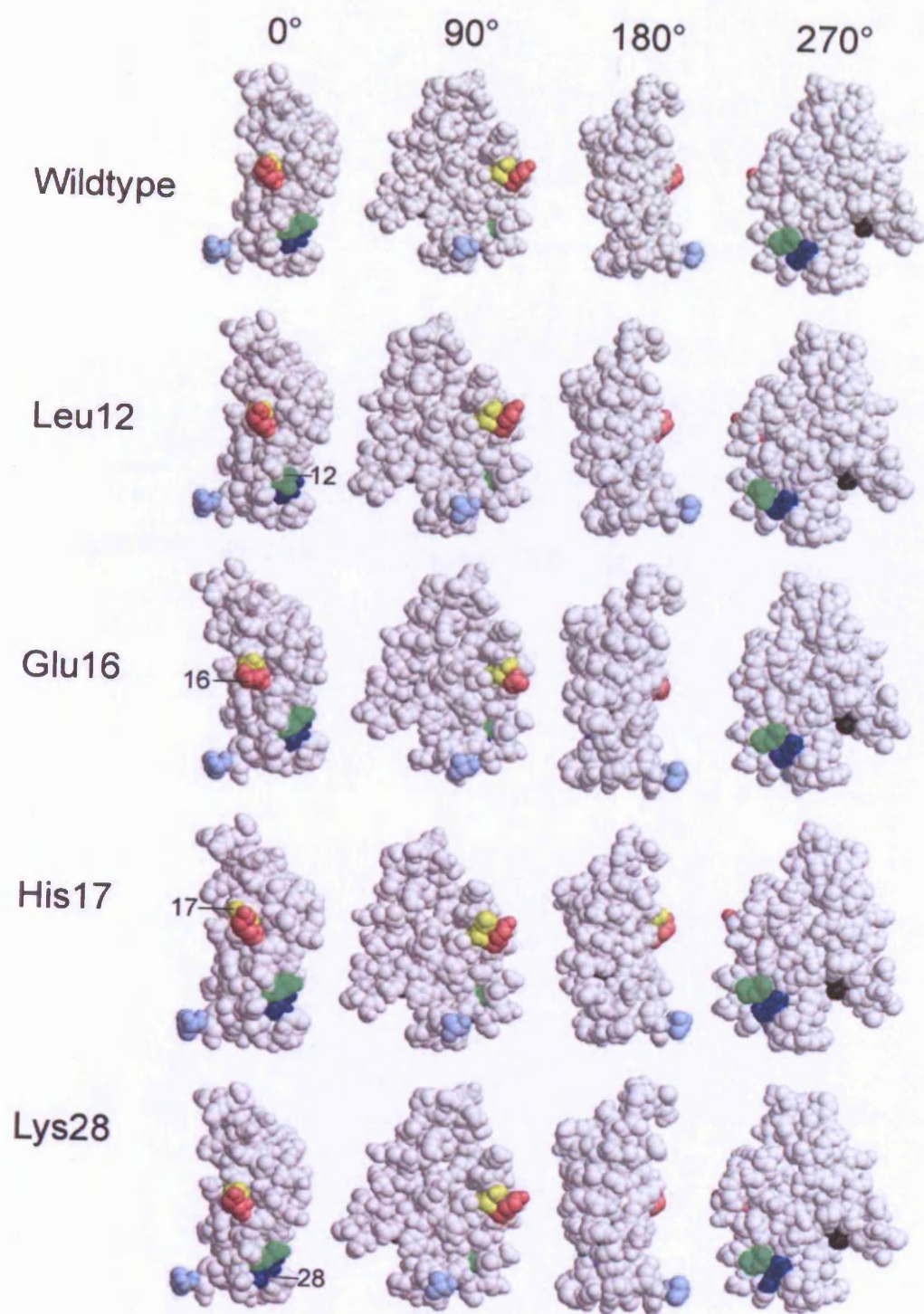


Figure 6.2: NuPAGE gel analysis of the expression time course for his-MSP1₁₉ variants in *Pichia pastoris*.

50 ml cultures of GS115 cells containing multiple copies of the wildtype and variant MSP1₁₉ genes were induced with methanol to express the proteins over 96 hours. 1 ml samples were removed every 24 hours. The supernatant was concentrated 10 times and run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer under reducing conditions and stained with Coomassie blue. The NuPAGE gels for the 24, 48, 72 and 96 hour time points are shown. Lane 1 – molecular mass markers, lane 2 – GS115 control cells, lane 3 – GS115 cells expressing wildtype MSP1₁₉, lane 4 – GS115 cells expressing Arg12→Leu MSP1₁₉ variant, lane 5 – GS115 cells expressing Lys16→Glu MSP1₁₉ variant, lane 6 – GS115 cells expressing Asn17→His MSP1₁₉ variant, lane 7 – GS115 cells expressing Glu28→Lys MSP1₁₉ variant, lane 8 – GS115 cells expressing double Lys16→Glu/Glu28→Lys MSP1₁₉ variant. The bands between 15 and 20 kDa are the his-MSP1₁₉ variant proteins.

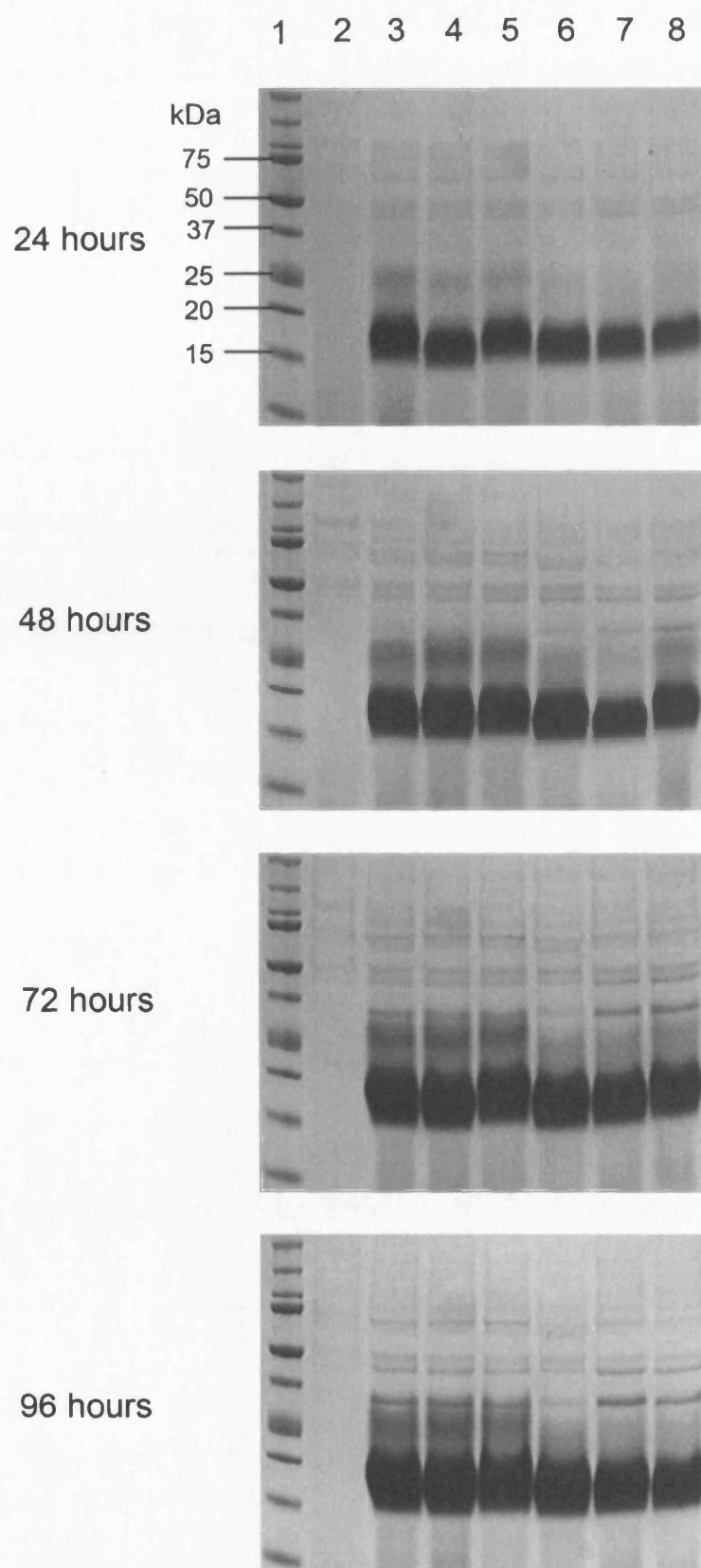


Figure 6.3: NuPAGE gel analysis of the purification of 50 ml cultures of wildtype and Glu28→Lys MSP1₁₉ variant his-tagged proteins expressed in *Pichia pastoris*.

The wildtype and Glu28→Lys MSP1₁₉ variant were expressed in 50 ml yeast cultures by methanol induction (1 %) for 96 hours. The protein was purified using a Ni-NTA agarose column washing with 10 mM imidazole, followed by 20 mM imidazole and eluting with 250 mM imidazole. The protein samples were run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer under reducing conditions and stained with Coomassie blue.

A: The NuPAGE gel for wildtype MSP1₁₉ purification. Lane 1 – molecular weight markers, lane 2 – culture supernatant before purification, lane 3 – flow through from Ni-NTA agarose column, lanes 4, 5 – column washes, lane 6 – elution, lane 7 – dialysed elution. The wildtype MSP1₁₉ protein runs around 15 kDa.

B: The NuPAGE gel for Glu28→Lys MSP1₁₉ variant purification. Lane 1 – molecular weight markers, lane 2 – culture supernatant before purification, lane 3 – flow through from Ni-NTA agarose column, lanes 4, 5 – column washes, lane 6 – elution, lane 7 – dialysed elution. The Glu28→Lys MSP1₁₉ variant protein runs around 15 kDa.

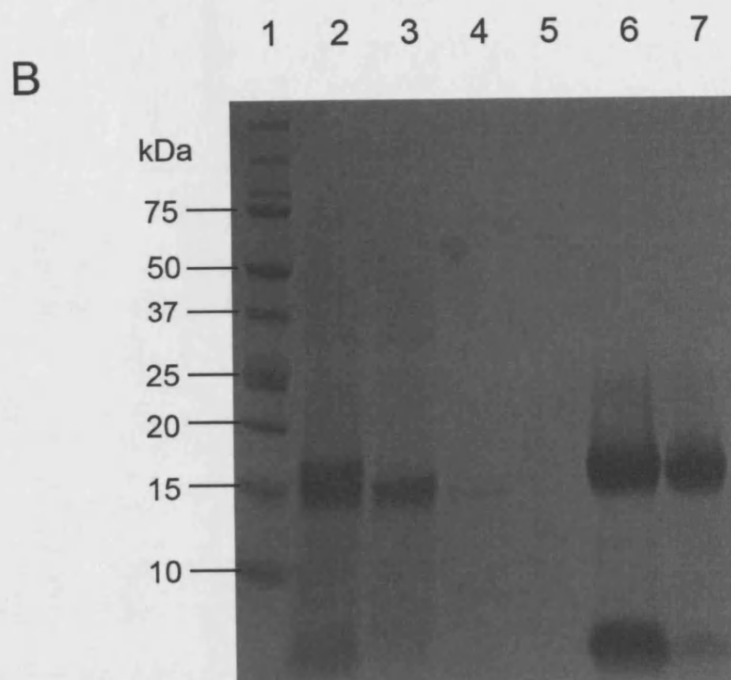
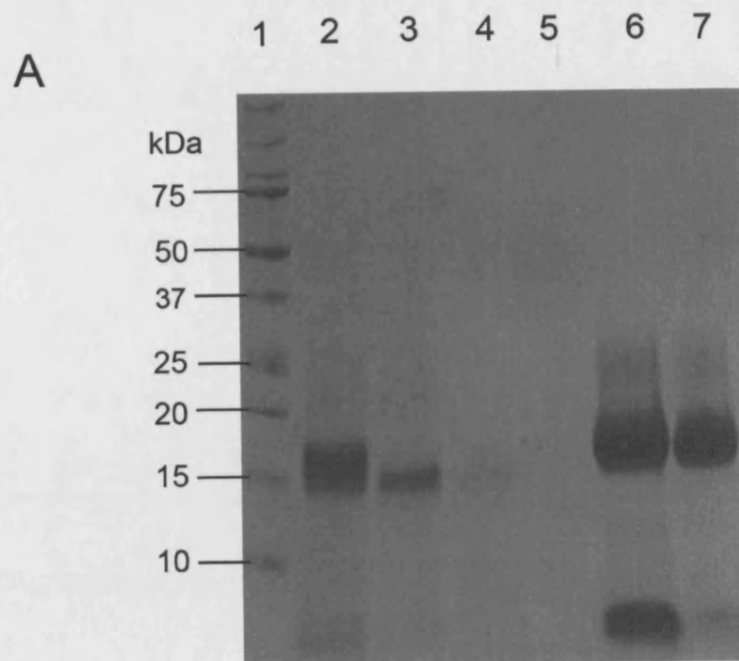


Figure 6.4: NuPAGE gel analysis of the purification of ^{15}N labelled wildtype and Glu28→Lys MSP1₁₉ variant his-tagged proteins for NMR analysis.

The wildtype and Glu28→Lys MSP1₁₉ variant were expressed in 4 × 500 ml yeast cultures by methanol induction (1 %) for 72 hours. The protein was purified using a batch method with Ni-NTA agarose washing with 10 mM imidazole, followed by 20 mM imidazole and eluting with 250 mM imidazole. The protein samples were run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer under reducing conditions and stained with Coomassie blue.

A: The NuPAGE gel for wildtype MSP1₁₉ purification. Lane 1 – molecular weight markers, lane 2 – culture supernatant before purification, lane 3 – flow through from Ni-NTA agarose column, lanes 4, 5 – column washes, lane 6 – elution, lane 7 – 5 µl dialysed elution, lane 8 – 10 µl dialysed elution, lane 9 – 15 µl dialysed elution. The wildtype MSP1₁₉ protein runs around 15 kDa.

B: The NuPAGE gel for Glu28→Lys MSP1₁₉ variant purification. Lane 1 – molecular weight markers, lane 2 – culture supernatant before purification, lane 3 – flow through from Ni-NTA agarose column, lanes 4, 5 – column washes, lane 6 – elution, lane 7 – 5 µl dialysed elution, lane 8 – 10 µl dialysed elution, lane 9 – 15 µl dialysed elution. The Glu28→Lys MSP1₁₉ variant protein runs around 15 kDa.

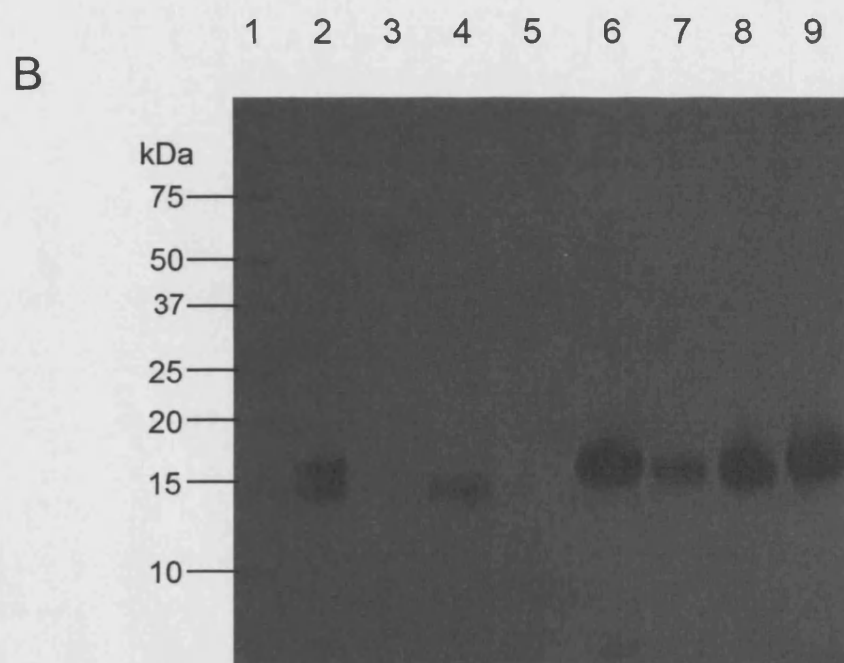
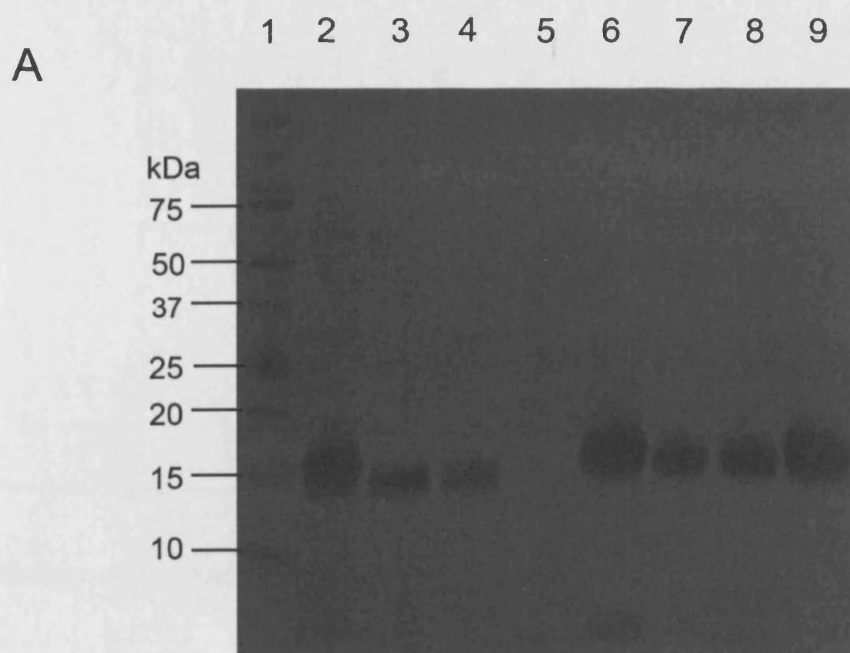


Figure 6.5: 1D ^1H NMR Spectra for wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ proteins.

A: 1D ^1H NMR Spectra for wildtype MSP1₁₉. The labels indicate the approximate chemical shift ranges for protons in proteins (Edwards & Reid, 2000).

B: 1D ^1H NMR Spectra for Glu28→Lys MSP1₁₉ variant protein.

The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

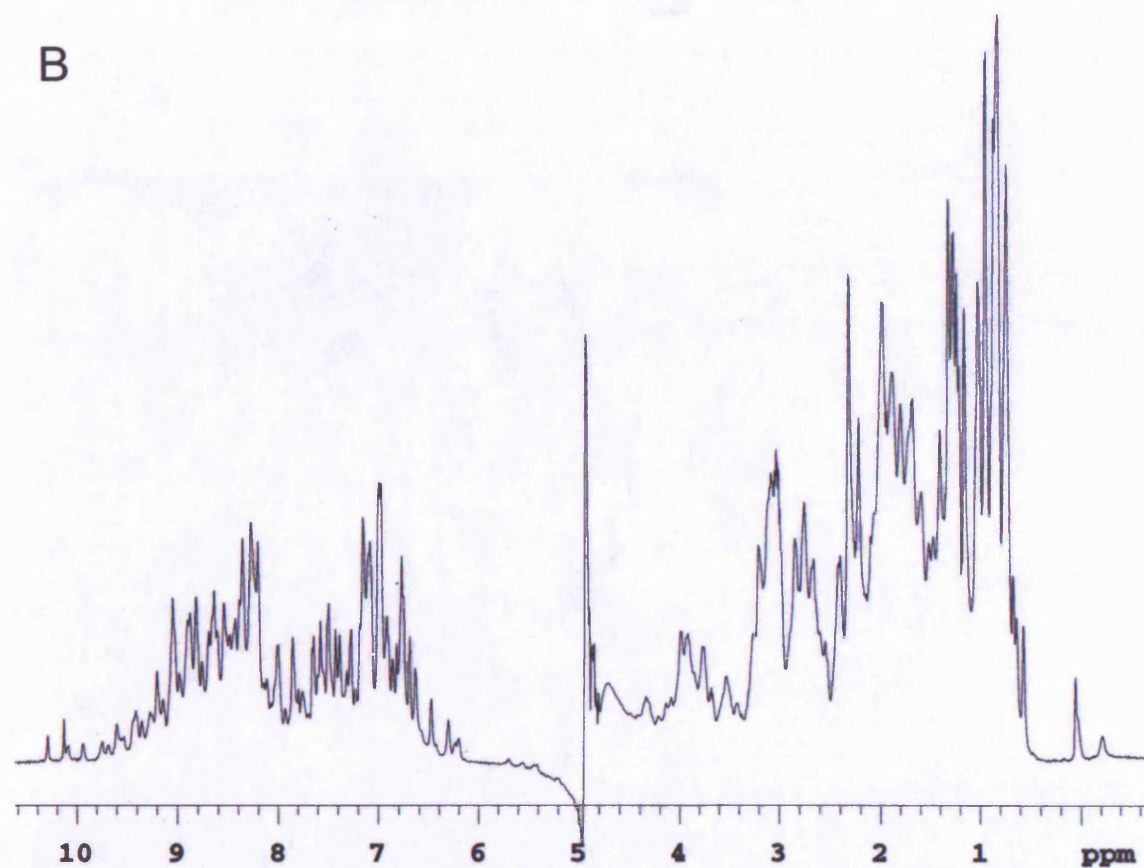
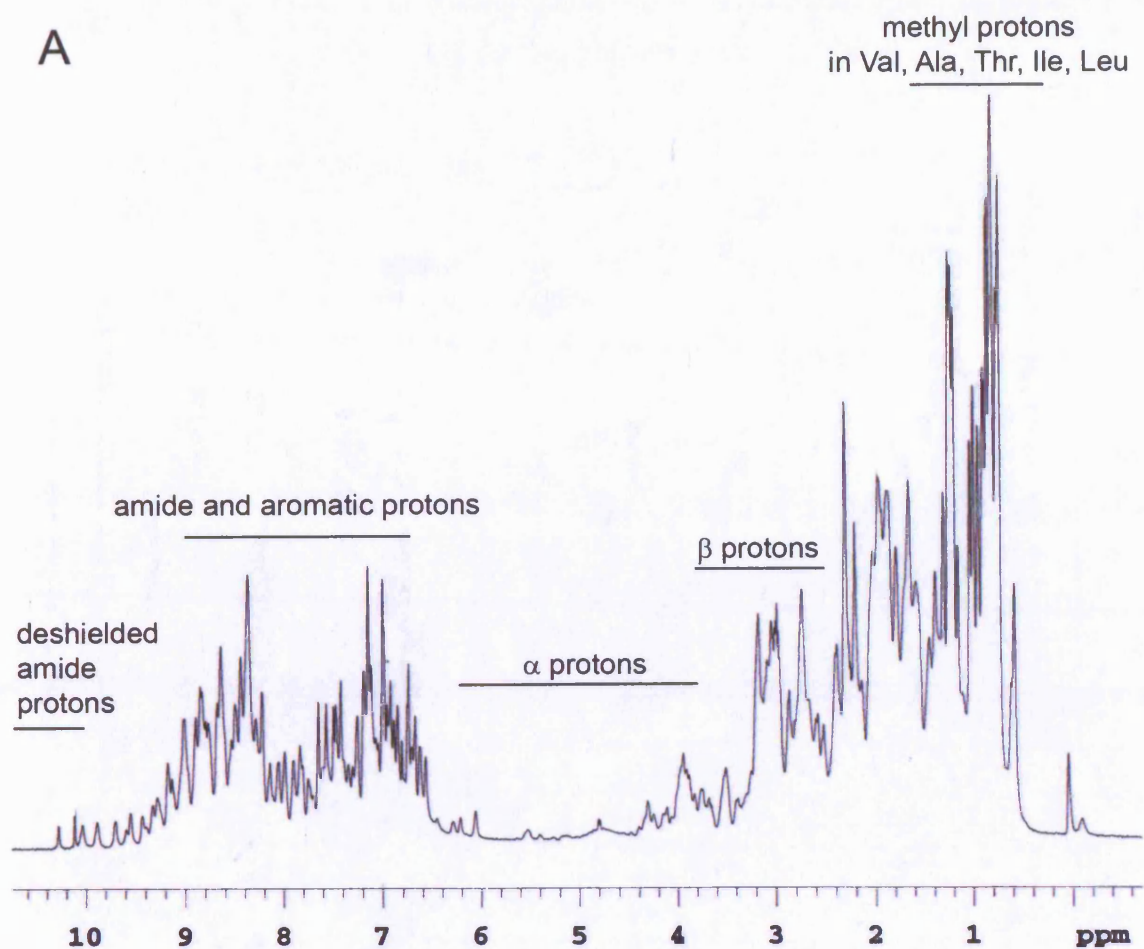


Figure 6.6: ^{15}N -HSQC Spectrum of wildtype MSP1₁₉.

A: Schematic representation of a ^{15}N -HSQC Spectrum.

B: ^{15}N -HSQC Spectrum of wildtype MSP1₁₉ with NH assignments. The spectrum was assigned using the data from HNCACB, CBCACONH and ^{15}N -NOESY-HSQC NMR spectra.

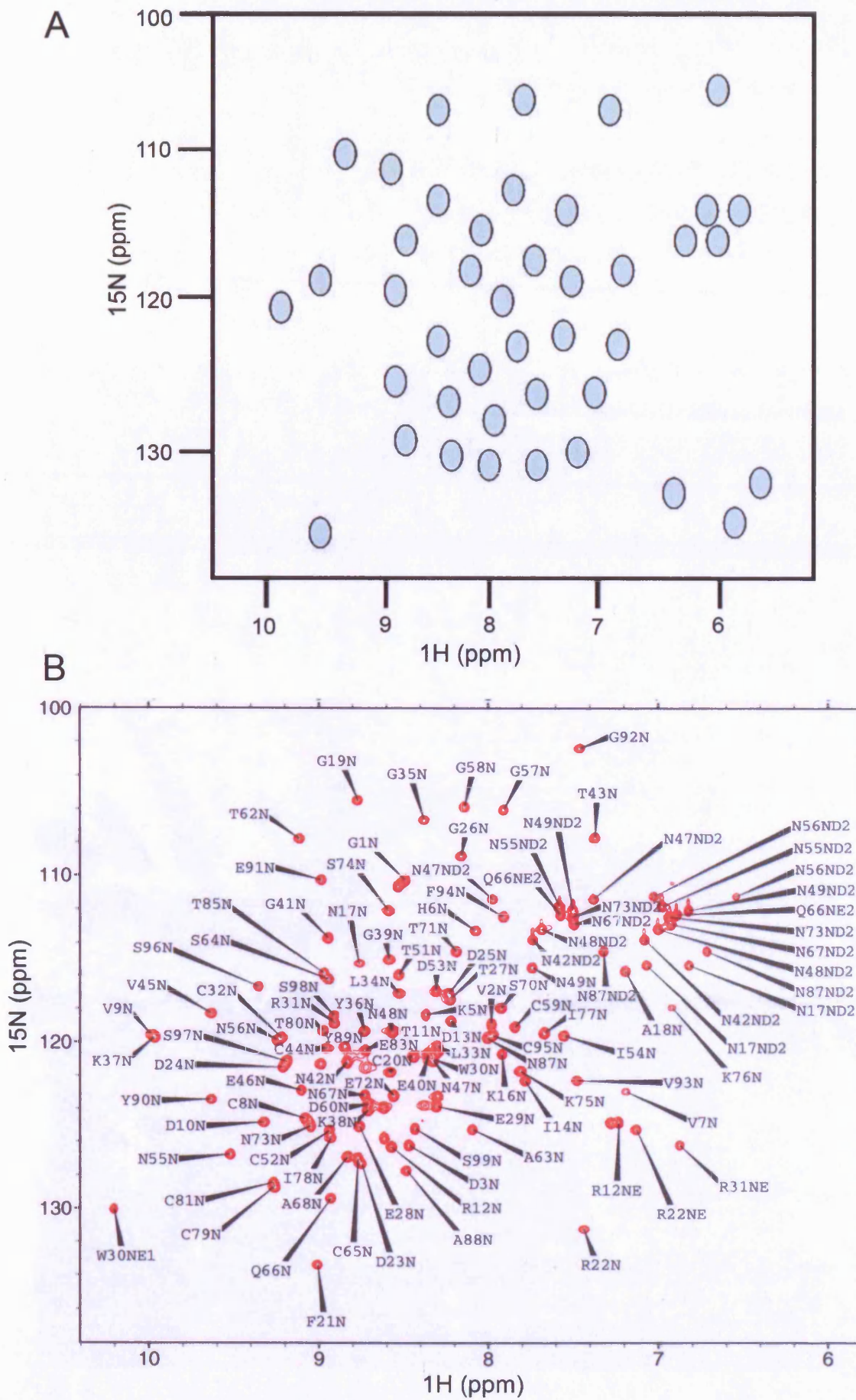


Figure 6.7: Comparison of Arg12→Leu MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of Arg12→Leu MSP1₁₉ variant is shown in black overlayed on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the Arg12→Leu MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum.

The following equation was used to combine the chemical shifts:

$$\text{Absolute } ((\text{variant } ^{15}\text{N chemical shift} - \text{wildtype } ^{15}\text{N chemical shift}) \div 5) + \text{Absolute } (\text{variant } ^1\text{H chemical shift} - \text{wildtype } ^1\text{H chemical shift}) \div 2$$

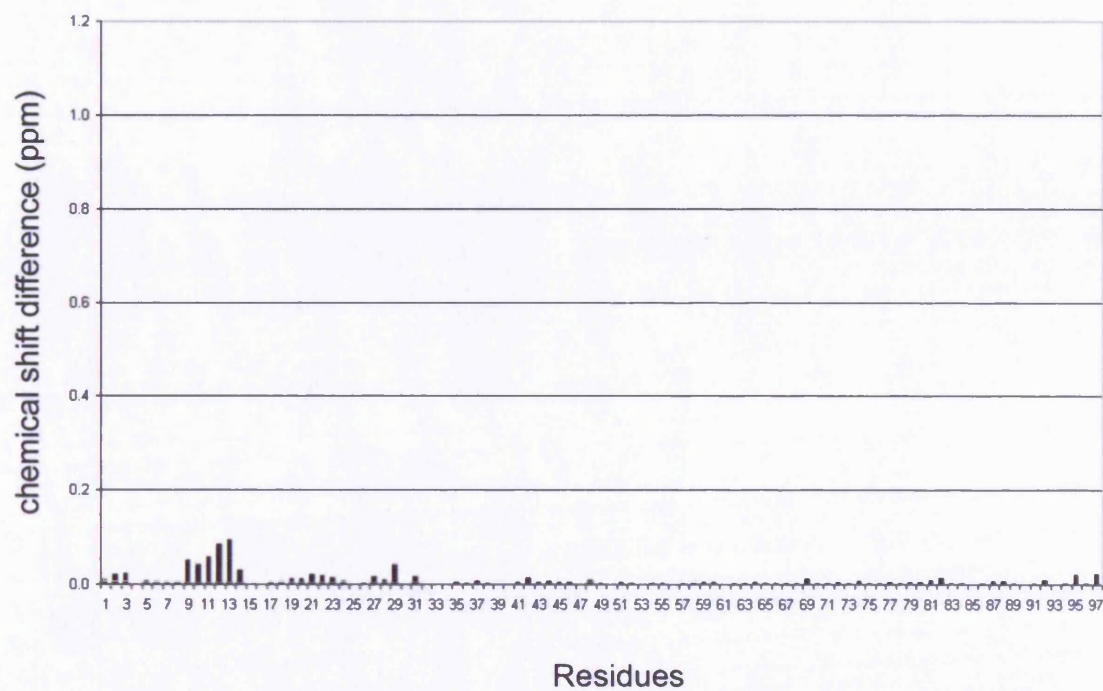
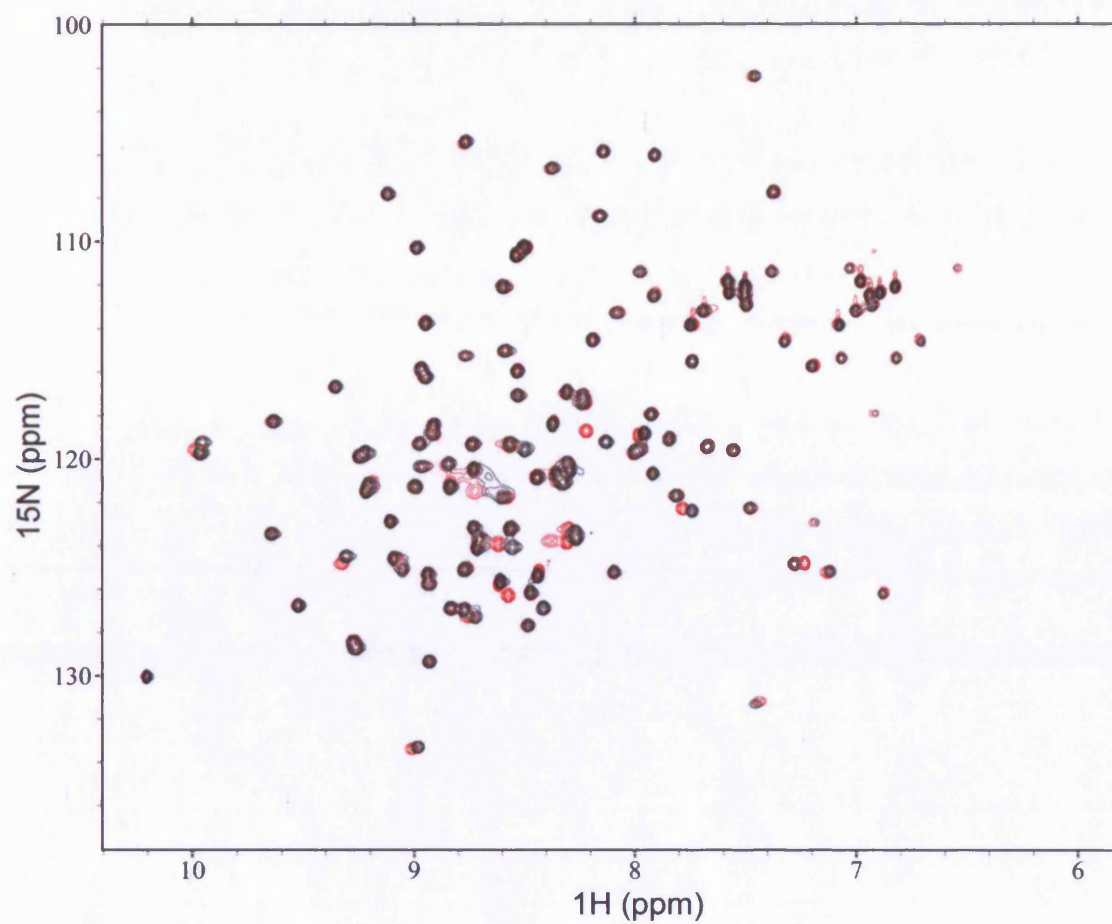


Figure 6.8: Comparison of Lys16→Glu MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of Lys16→Glu MSP1₁₉ variant is shown in cyan overlayed on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the Lys16→Glu MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum.

The following equation was used to combine the chemical shifts:

Absolute ((variant ¹⁵N chemical shift – wildtype ¹⁵N chemical shift) ÷ 5) + Absolute (variant ¹H chemical shift – wildtype ¹H chemical shift) ÷ 2

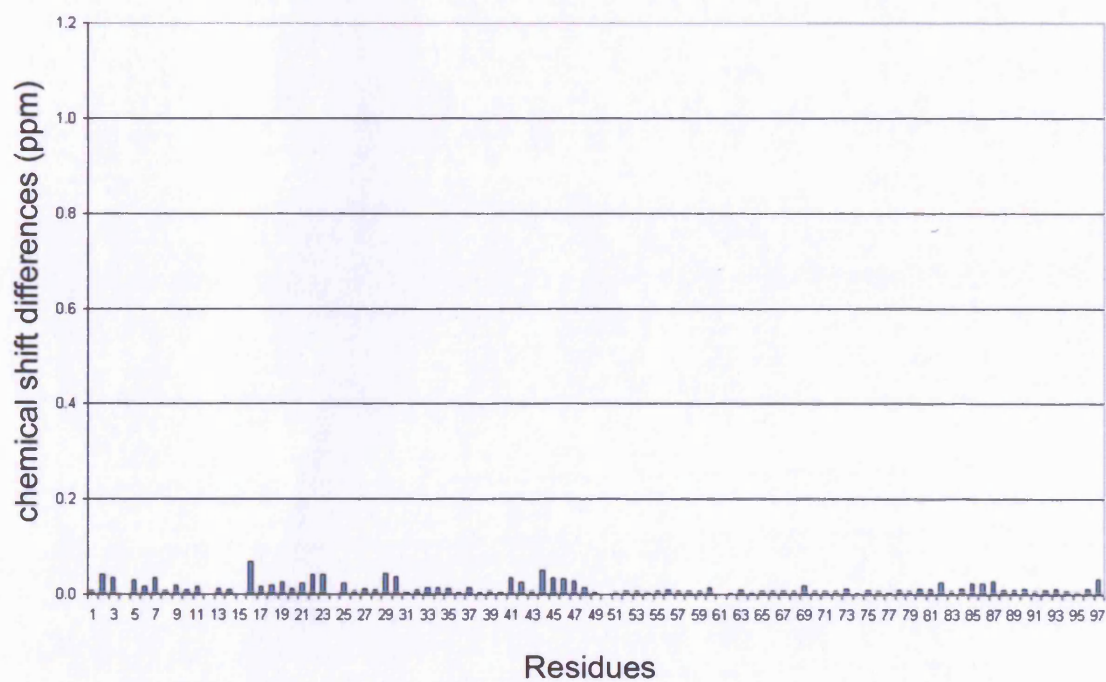
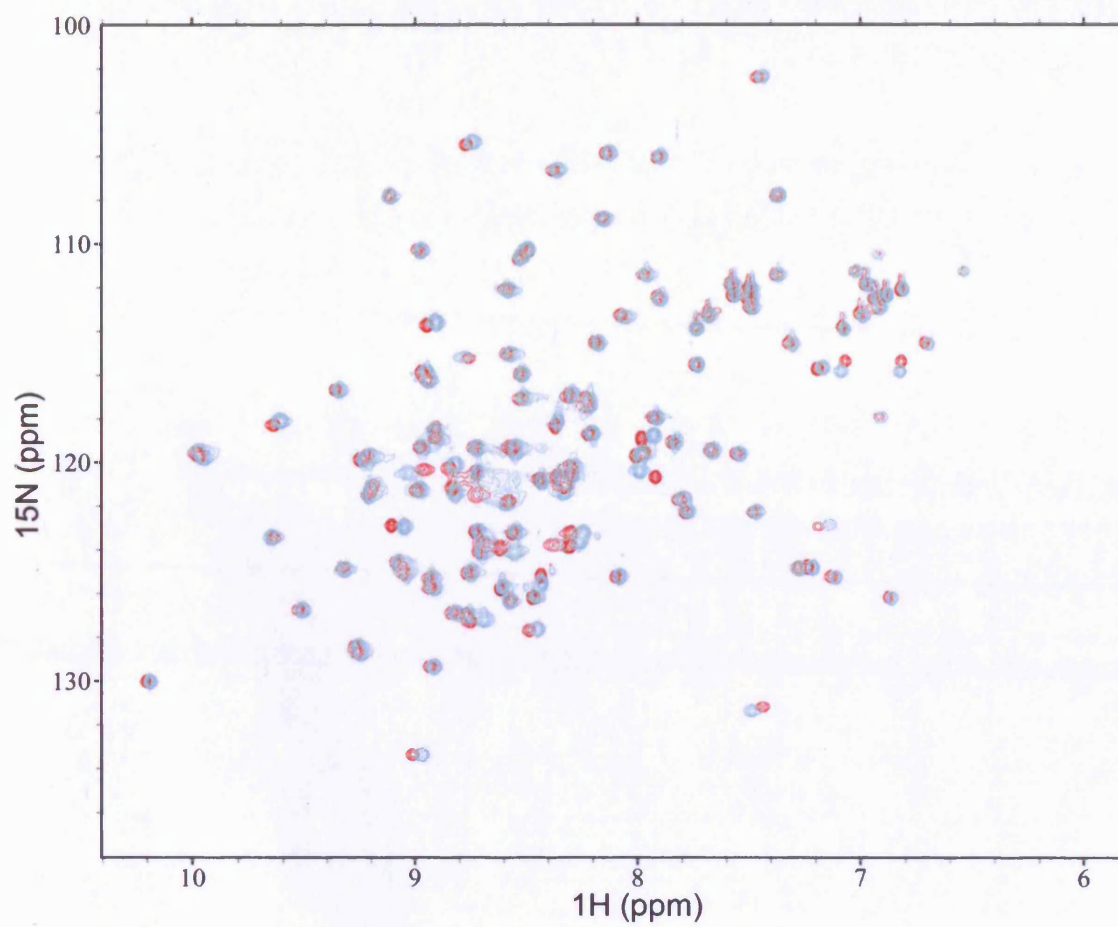


Figure 6.9: Comparison of Asn17→His MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of Asn17→His MSP1₁₉ variant is shown in purple overlayed on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the Asn17→His MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum.

The following equation was used to combine the chemical shifts:

Absolute ((variant ¹⁵N chemical shift – wildtype ¹⁵N chemical shift) ÷ 5) + Absolute (variant ¹H chemical shift – wildtype ¹H chemical shift) ÷ 2

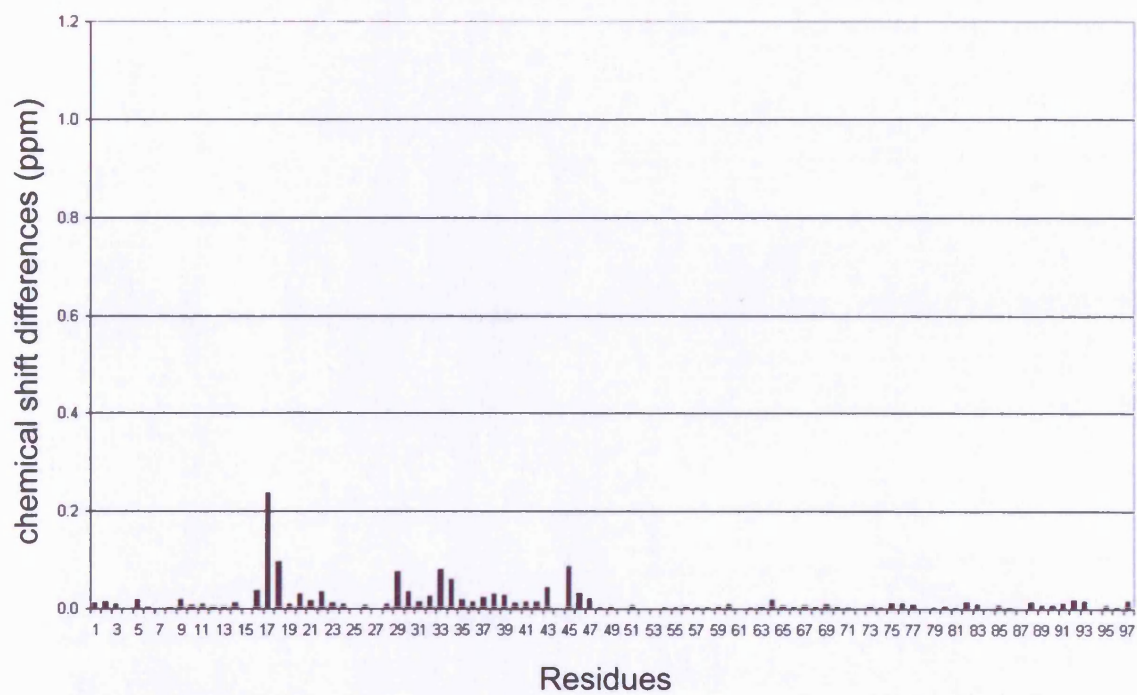
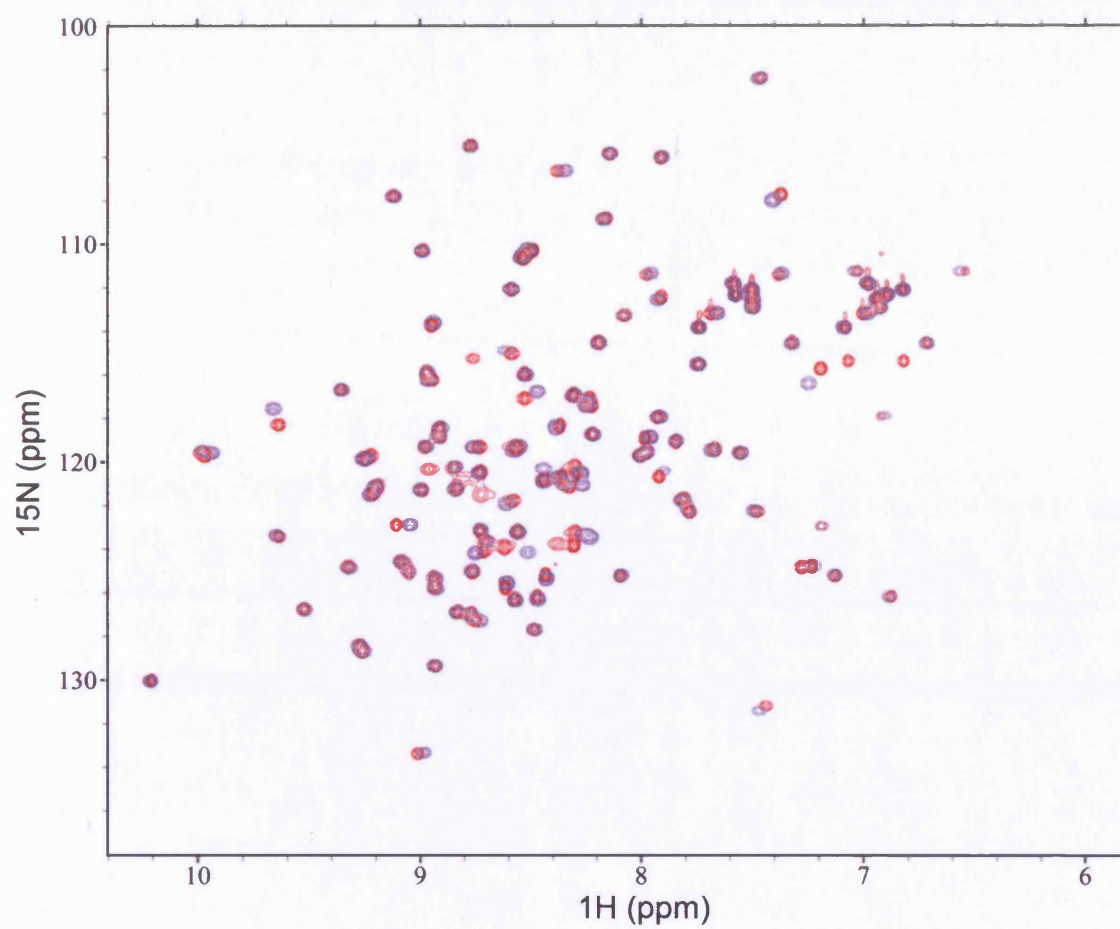


Figure 6.10: Comparison of Glu28→Lys MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of Glu28→Lys MSP1₁₉ variant is shown in blue overlaid on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the Glu28→Lys MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum.

The following equation was used to combine the chemical shifts:

Absolute ((variant ¹⁵N chemical shift – wildtype ¹⁵N chemical shift) ÷ 5) + Absolute (variant ¹H chemical shift – wildtype ¹H chemical shift) ÷ 2

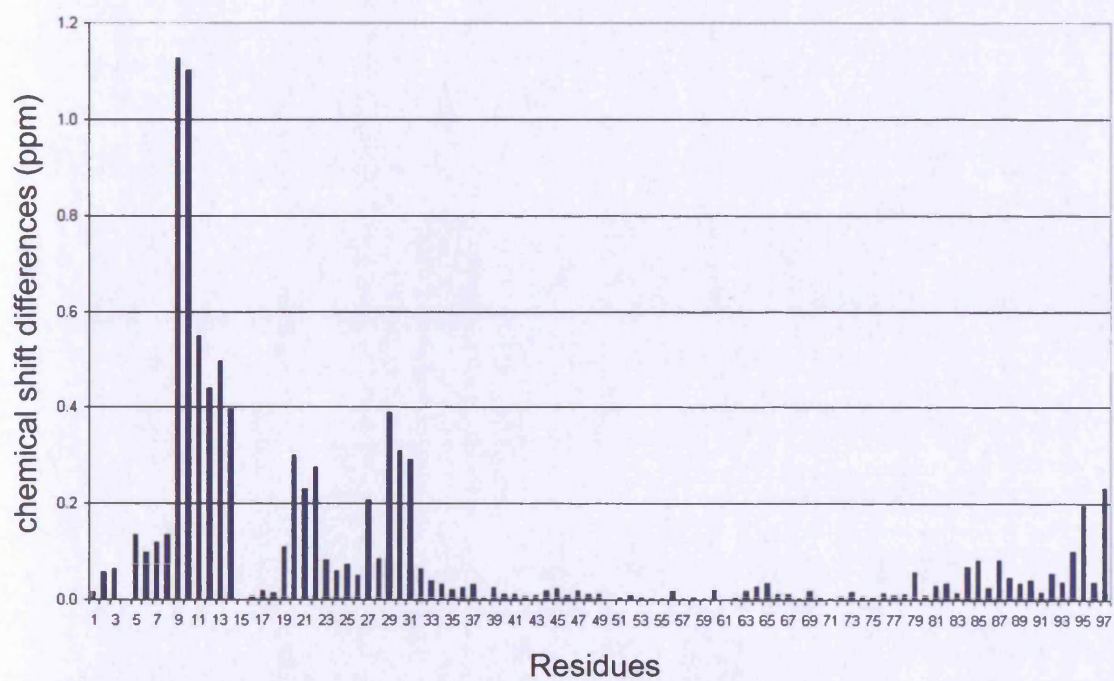
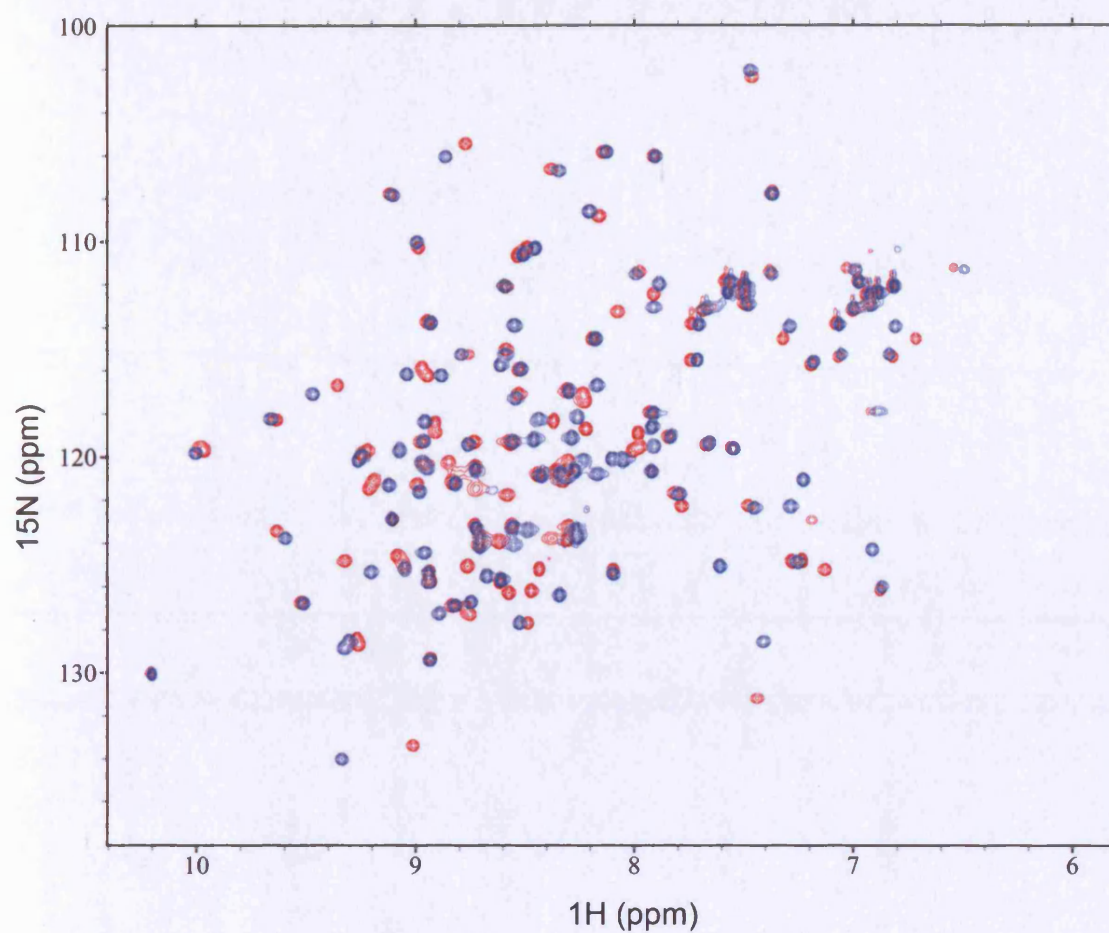


Figure 6.11: Comparison of double Lys16→Glu/Glu28→Lys MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in gold overlayed on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum. The following equation was used to combine the chemical shifts:
Absolute ((variant ¹⁵N chemical shift – wildtype ¹⁵N chemical shift) ÷ 5) + Absolute (variant ¹H chemical shift – wildtype ¹H chemical shift) ÷ 2

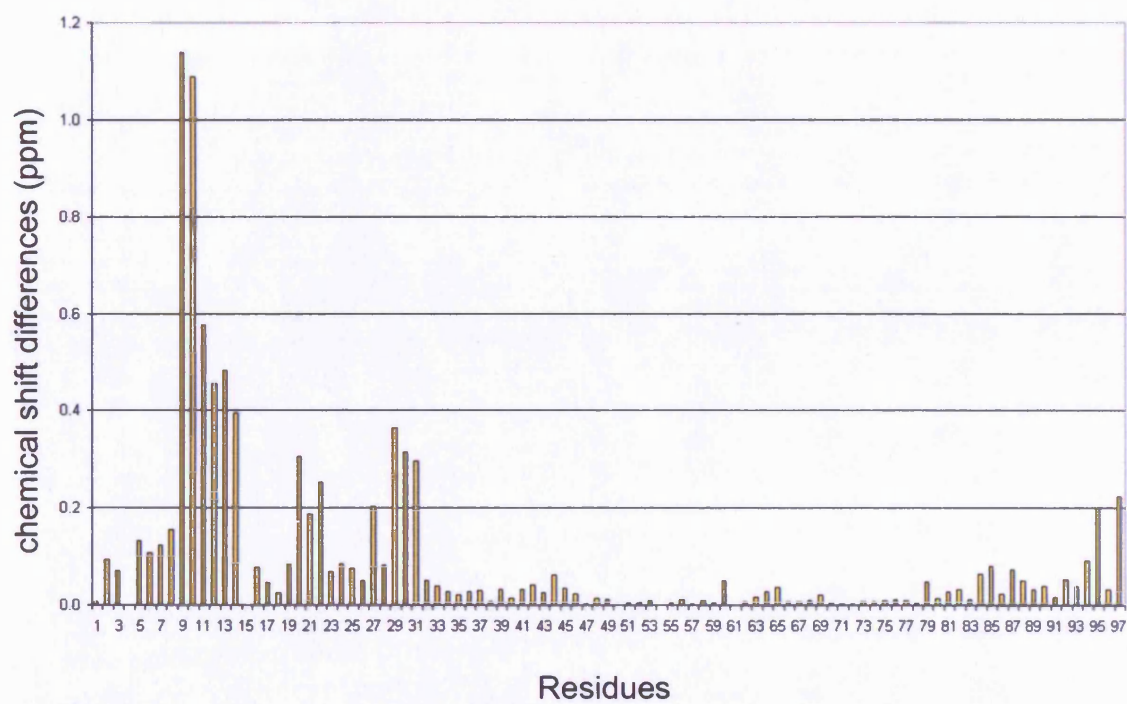
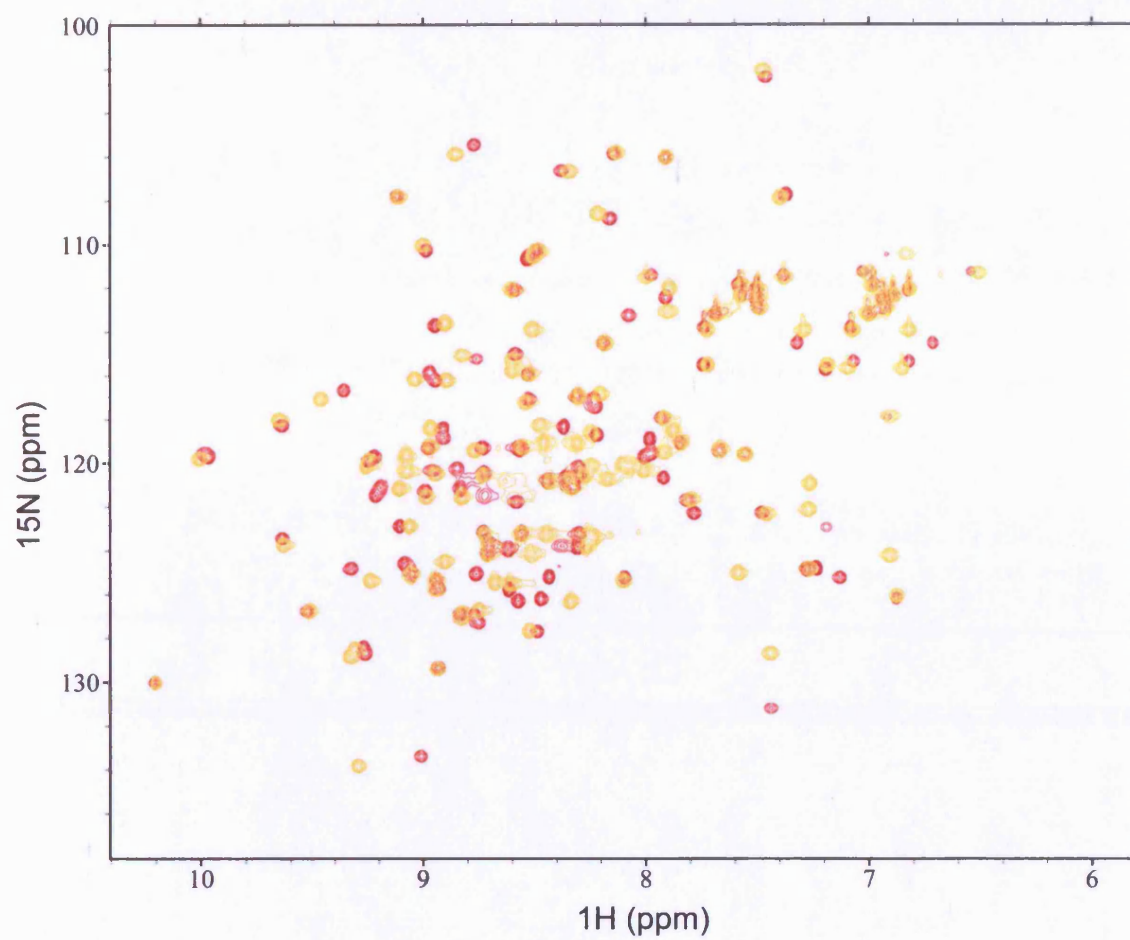


Figure 6.12: Comparison of double Lys16→Glu/Glu28→Lys and Glu28→Lys MSP1₁₉ variants ¹⁵N-HSQC NMR spectra.

The ¹⁵N-HSQC spectrum of double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in gold overlayed on top of the Glu28→Lys MSP1₁₉ variant spectrum (shown in blue) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

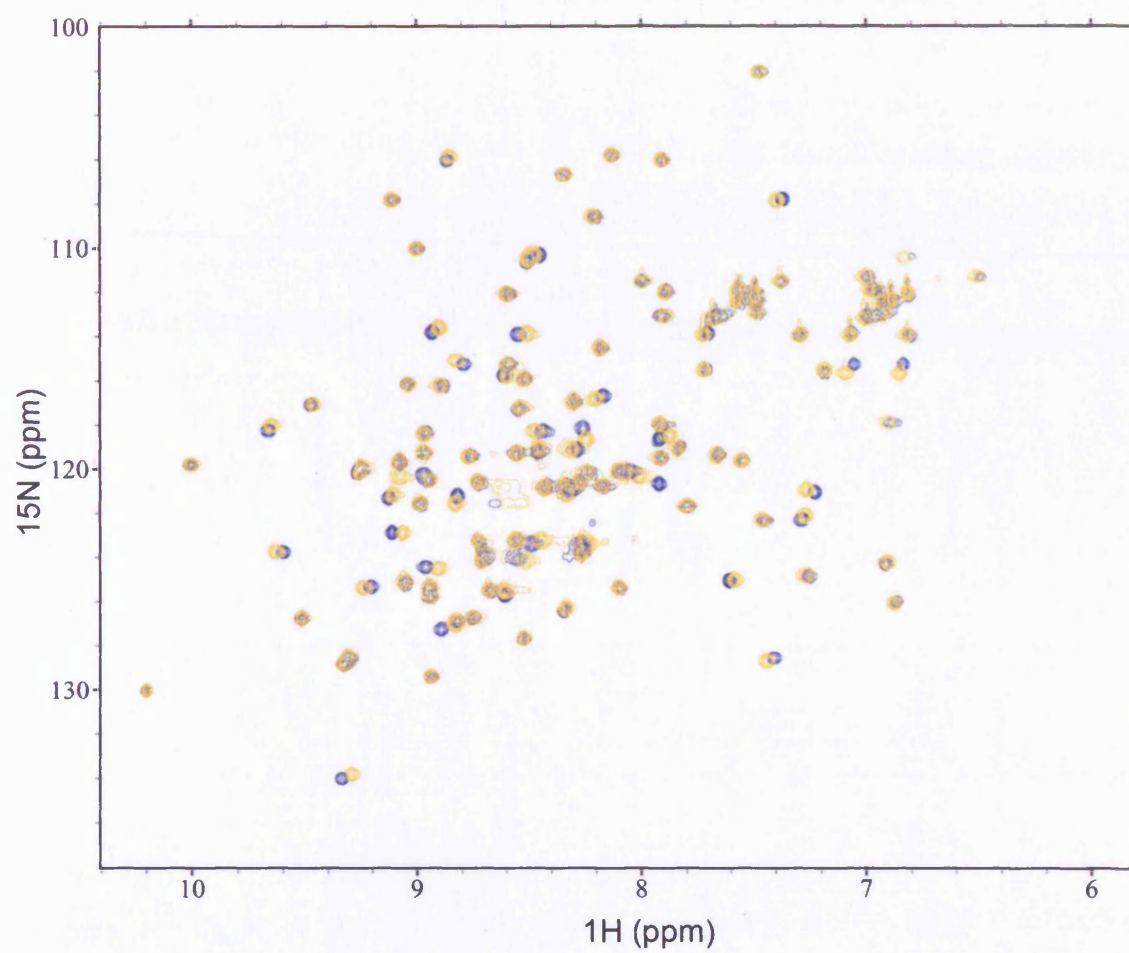


Figure 6.13: Comparison of the combined ^{15}N and ^1H chemical shift differences of double Lys16→Glu/Glu28→Lys, Glu28→Lys and Lys16→Glu MSP1₁₉ variants compared to wildtype MSP1₁₉.

Histogram showing the combined ^{15}N and ^1H chemical shift differences for the double Lys16→Glu/Glu28→Lys, Glu28→Lys and Lys16→Glu MSP1₁₉ variants ^{15}N -HSQC spectra compared to the wildtype spectrum. The double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in gold. The Glu28→Lys MSP1₁₉ variant is shown in blue and the Lys16→Glu MSP1₁₉ variant is shown in cyan.

The following equation was used to combine the chemical shifts:

$$\text{Absolute } ((\text{variant } ^{15}\text{N chemical shift} - \text{wildtype } ^{15}\text{N chemical shift}) \div 5) + \text{Absolute } ((\text{variant } ^1\text{H chemical shift} - \text{wildtype } ^1\text{H chemical shift}) \div 2)$$

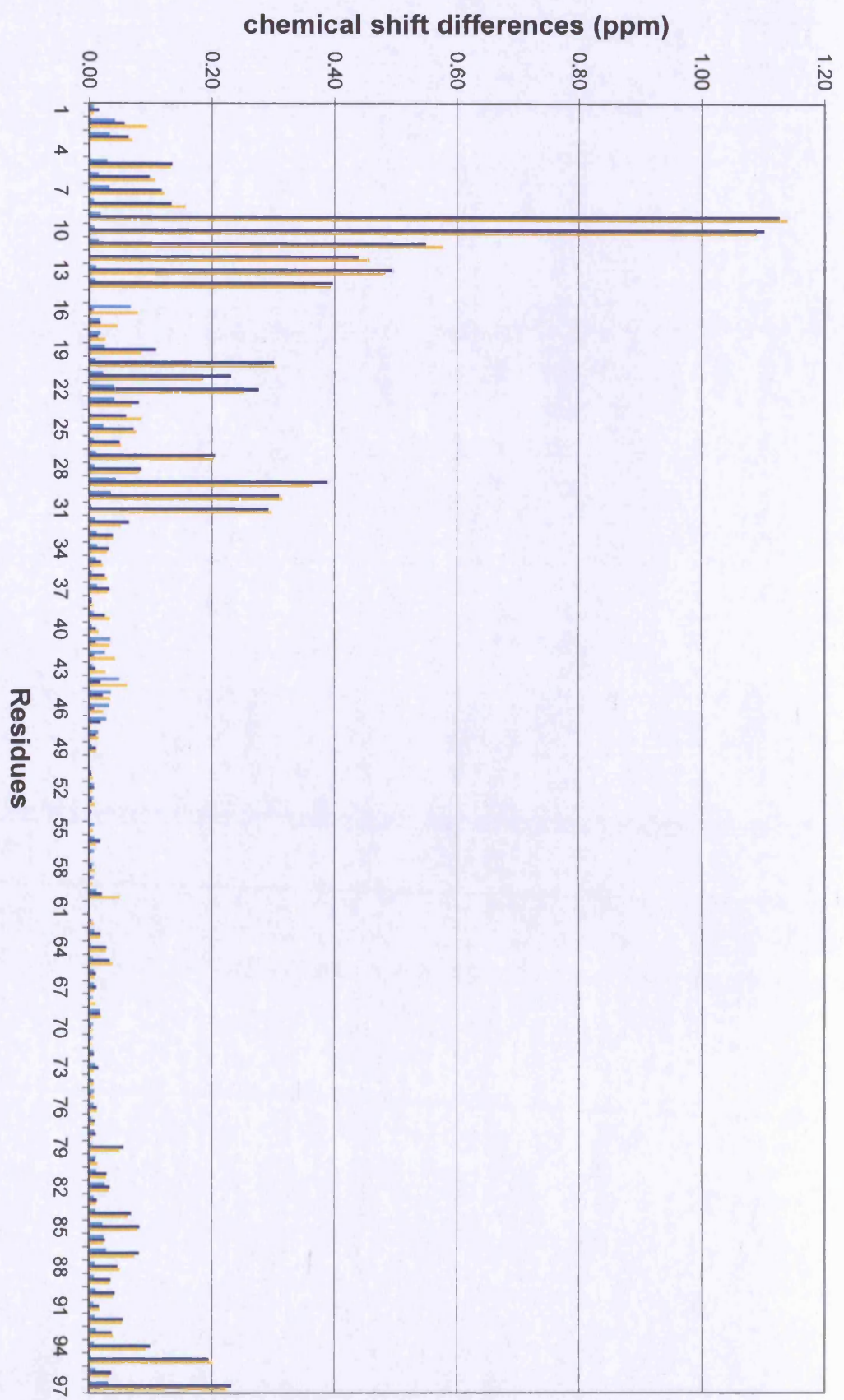
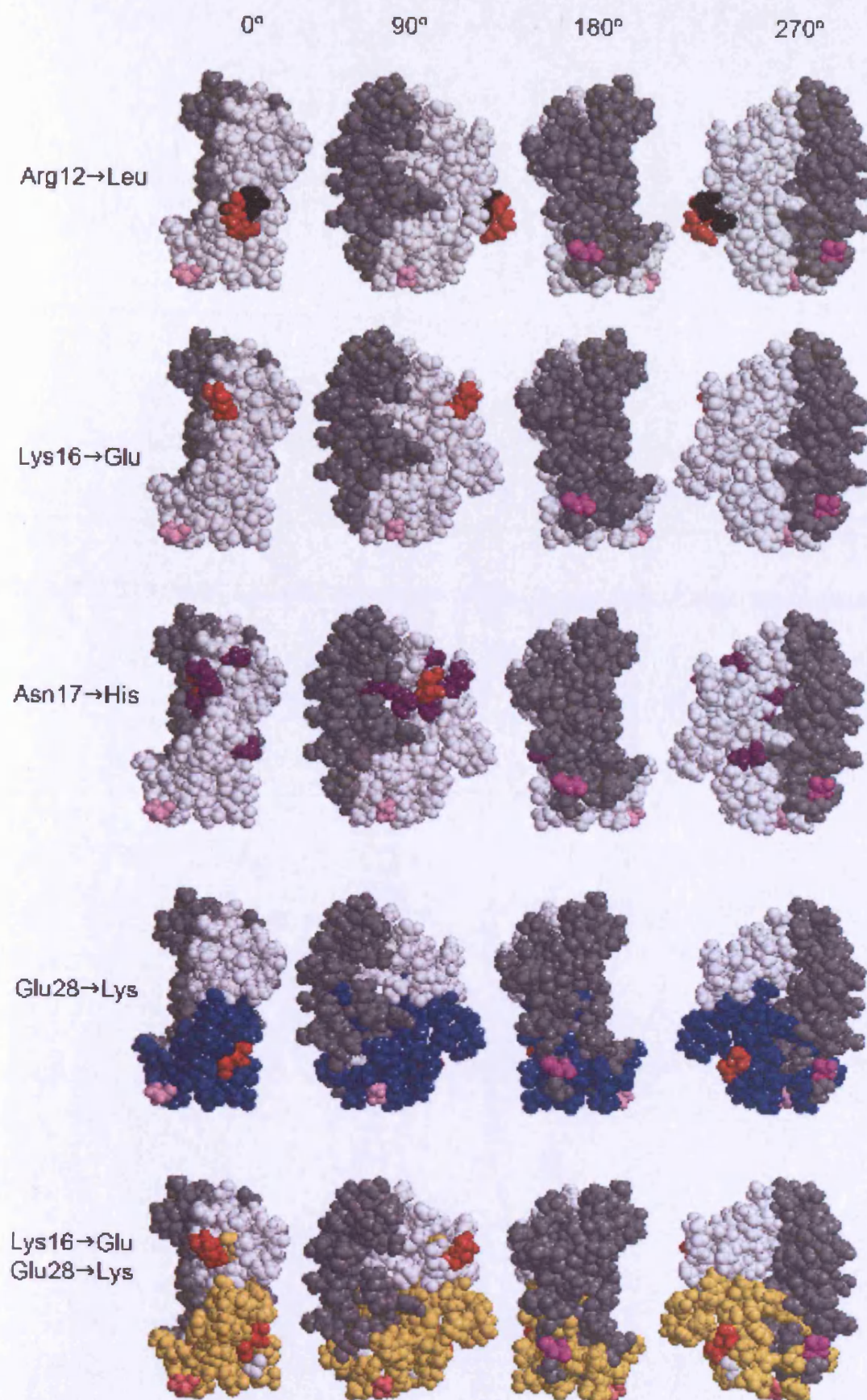


Figure 6.14: Mapping the NH peaks that have moved in the ^{15}N -HSQC spectra onto the 3D structure of *P. yoelii* MSP1₁₉.

The residues corresponding to the NH peaks that have moved more than 0.05 ppm for the combined chemical shift difference in the ^{15}N -HSQC spectra for the variants compared to the wildtype MSP1₁₉ have been mapped onto the best energy NMR structure for the wildtype *P. yoelii* MSP1₁₉. The calculation of this structure is discussed in chapter 9. The residues that have been changed in the variants are shown in red. The residues corresponding to the NH peaks that have moved more than 0.05 ppm are coloured as follows: Arg12→Leu in black; Lys16→Glu in cyan; Asn17→His in purple; Glu28→Lys in blue; the double Lys16→Glu/Glu28→Lys in gold. The first EGF domain is shown in white and the second EGF domain is shown in grey. The C-terminal residue is shown in bright pink and the N-terminal residue is shown in light pink. For the Lys16→Glu variant there are no amino acids coloured cyan as the chemical shifts changes are all less than 0.05 ppm.



6.5 Discussion

The *in silico* variations of the residues on the *P. yoelii* model showed only minor visible structural changes for the Lys16→Glu, Asn17→His and Glu28→Lys variants. For Arg12→Leu variants there were no visible structural changes to the protein. This was a model structure and may not represent what would be seen in the actual protein.

The data has shown that *P. pastoris* is a good system to create milligram quantities of ¹⁵N labelled protein and that by producing the protein in a secreted form the protein could be purified for NMR analysis.

The spread of NH peaks in the wildtype MSP1₁₉ spectra has confirmed that the protein is folded because the glycine peaks at the top of the spectra are clearly spread out which would indicate they were in a folded protein. If all the glycines, for example, had appeared as a single peak or very close together this could have suggested the protein was not folded. The overall dispersion of the peaks in the spectra for all the MSP1₁₉ variants was very similar to the wildtype protein. This suggests that the amino acid variations to the proteins had not completely denatured the proteins or affected the overall folding of the proteins. The spectra for the Arg12→Leu variant and Lys16→Glu variant showed that only the NH peaks corresponding to amino acids in direct contact with residues 12 or 16 had moved. This indicates that any changes in the protein structure as a result of the residues 12 or 16 variations were localised around the individual residues. This would suggest that there were no significant structural differences between these proteins and the wildtype which means that I can be confident that the changes in antibody binding seen in chapter 3 are a direct result of the changes to the individual residues. For the residue 12 and 16 variations this agrees with the prediction of the affect of changing the amino acids from the *in silico* variations on the *P. yoelii* model. The spectra for the Asn17→His variant showed that NH peaks corresponding to residues in close proximity to residue 17 had moved and not just those residues in direct contact with residue 17. This may be because the amino acid was changed to a histidine. Histidine is an aromatic amino acid and the ring produces a ring current effect. This means that the protons in the plane of the ring can be shifted to a lower field and the protons that are above and below the plane of the ring can be shifted

to a higher field and the shift can be large. This can mean that the affect of inserting an aromatic amino acid into the protein can have a large affect on the chemical shifts of the NMR peaks but not necessarily such a large affect on the actual protein structure. There is movement to residue 29 which is further away from residue 17. The movement to this residue may be due to an error in the assignment of this peak in the Asn17→His variant spectrum as this is an area where there is overlap of NH peaks in the wildtype spectra which could result in difficulties in confirming where the NH peak in Asn17→His has moved to. In order to confirm if residue 29 has moved, 3D NMR studies would need to be carried out to confirm where the peak for residue 29 is in the Asn17→His variant. The NMR data for Asn17→His suggests that the changes in the antibody binding (seen in chapter 3) were a direct result of the variation to residue 17 as most of the NH peaks that have moved corresponded to amino acids in close proximity to residue 17. This is in agreement with the prediction for the residue 17 change on the *P. yoelii* model as it showed only minor visible structural differences.

The Glu28→Lys variant HSQC spectrum was visually very different from the wildtype. Most of the NH peaks corresponding to the first EGF domain had moved in the spectra and some by a large amount. For example, valine 9 has shifted 1.49 ppm and aspartic acid 10 has shifted 1.27 ppm upfield in the ¹H dimension suggesting that there may be hydrogen bonds missing or broken involving these NHs in Glu28→Lys variant. This suggests that there is a lot of structural perturbation in the first EGF domain as a result of the residue 28 variation. There were also NH peaks corresponding to the end of the second EGF domain that had moved in the spectra. This indicates that the residue 28 variation has had a large affect on the structure as it has affected the second domain when residue 28 is found in the first EGF domain. The Glu28→Lys variant spectrum data therefore indicates that I cannot be confident that the changes in antibody binding and loss of protection *in vivo* for Glu28→Lys variant were a direct result of the change to residue 28. The changes in antibody binding and *in vivo* loss of protection are therefore a result of the structural changes that have occurred due to the change in residue 28. The ¹⁵N-HSQC spectroscopy alone can only identify the areas of the protein that may have changed it cannot confirm how those areas have changed. The large change in the structure of Glu28→Lys variant was not predicted by the *in silico* variation of this residue in the model. In the model, the only change that was seen was a

difference in the angle of residue 28 relative to residue 12. The difference between the *in silico* prediction and the NMR prediction may be because for the *in silico* prediction the amino acid has been inserted into the position for the residue and the optimum rotamer confirmation has been found. The *in silico* prediction has not moved the backbone and neighbouring amino acids as a result of the change in residue. This shows that *in silico* predictions of amino acid changes using Deepview/Swiss Pdb viewer cannot always give reliable predictions if the amino acid change leads to lots of changes to the surrounding amino acids in the protein. The spectrum for the double Lys16→Glu/Glu28→Lys variant was very similar to the spectrum for the Glu28→Lys variant and that the chemical shift differences were a result of the combination of the difference for the Lys16→Glu and Glu28→Lys variant. This shows that the antibody binding data to the double Lys16→Glu/Glu28→Lys variant was a result of the structural changes that had occurred due to the changes to residue 16 and 28 and not a direct affect of binding to residue 16 and 28.

The data presented in this chapter has indicated that the residue 28 variation causes significant structural perturbation to the protein and that the *in silico* variation of the protein model is unable to predict this. The data shows that further structural analysis is required to determine the affects on the structure of the residue 28 variation and to determine if this residue has an important function in the structure of the protein.

Chapter 7: Designing and analysing a Glu28→Gln MSP1₁₉ variant

7.1 Introduction

The antibody binding studies of residue 12, 16, 17 and 28 shown in chapter 3 indicated that all of the residues were important for antibody binding. Only one of the amino acid variants, Glu28→Lys showed an affect on protection obtained by immunisation with the protein *in vivo*. The NMR studies in chapter 6 showed however that the Glu28→Lys MSP1₁₉ variant displayed a high level of structural perturbation in comparison to the wildtype protein. This meant that I could not be confident that the effect on antibody binding and on *in vivo* protection was a result of the change to residue 28 and not the structural changes arising as a result of the change. In this chapter, I will describe the production and analysis of a Glu28→Gln MSP1₁₉ variant to examine whether or not residue 28 is important for antibody binding. Figure 7.1 shows the differences between glutamic acid, lysine and glutamine. The residue 28 change from glutamic acid to lysine involves a change in charge from negative to positive and an increase in size. If the glutamic acid residue was involved in a charged interaction with surrounding amino acids, this reversal in charge could have resulted in the amino acid repelling neighbouring residues resulting in a structural change. A change from glutamic acid to glutamine involves a change in charge from negative to no charge but no change in size. It could be predicted that if the glutamic acid residue was involved in a charge interaction that the loss of charge may stop this interaction but would not repel the surrounding amino acids. This could potentially lead to less of a structural change. In this chapter, I will also compare the structure of the Glu28→Gln MSP1₁₉ variant to the wildtype protein and Glu28→Lys MSP1₁₉ variant to examine whether this smaller difference has a less significant affect on the structure.

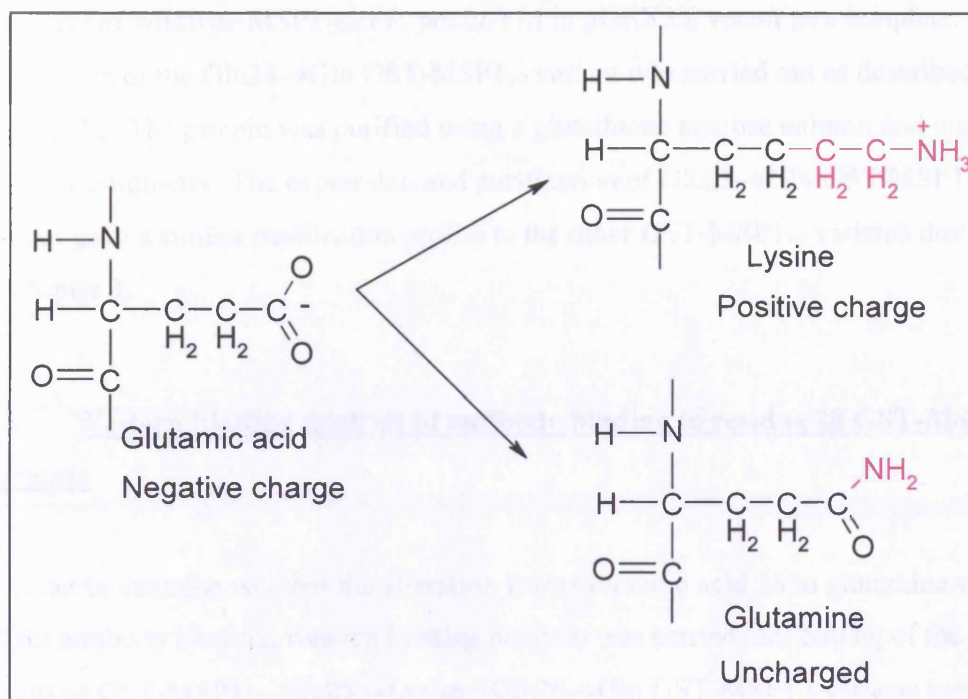


Figure 7.1: Differences between glutamic acid, lysine and glutamine side chains.

The left hand side shows the structure of glutamic acid 28 in the wildtype protein. The right hand side shows the structures of lysine and glutamine. The areas of the side chains that are different from glutamic acid are highlighted in pink. This figure was produced using MDL ISIS/Draw 2.5.

7.2 **Expression and purification of Glu28→Gln GST-MSP1₁₉ variant**

A Glu28→Gln GST-MSP1₁₉ variant was created by site-directed mutagenesis using the sequence of wildtype MSP1₁₉ of *P. yoelii* YM in pGEX3X vector as a template.

Expression of the Glu28→Gln GST-MSP1₁₉ variant was carried out as described section 2.2. The protein was purified using a glutathione agarose column and quantified using densitometry. The expression and purification of Glu28→Gln GST-MSP1₁₉ variant gave a similar purification profile to the other GST-MSP1₁₉ variants discussed in chapter 3.

7.3 **Western blotting analysis of antibody binding to residue 28 GST-MSP1₁₉ variants**

In order to examine whether the alteration from glutamic acid 28 to glutamine would affect antibody binding, western blotting analysis was carried out. 500 ng of the wildtype GST-MSP1₁₉, Glu28→Lys and Glu28→Gln GST-MSP1₁₉ variants were run on NuPAGE gels under non-reducing conditions for monoclonal antibody westerns and reducing conditions for anti-GST antibody westerns. A control western blot with anti-GST antibody was carried out as shown in figure 7.2 (top panel). This western blot shows binding to the wildtype and residue 28 GST-MSP1₁₉ variants confirming that the Glu28→Gln has been purified and quantified in the same way as the previously produced wildtype and Glu28→Lys GST-MSP1₁₉ variant. Western blotting analysis with B6, F5 and B10 antibodies (Spencer Valero *et al.*, 1998) show that the Glu28→Gln variation has the same affect on binding as the Glu28→Lys variation. The results are summarised in table 7.1.

Table 7.1: Effects of residue 28 variations on B6, F5 and B10 binding as shown by western blotting (figure 7.2)

Variant	B6	F5	B10
Glu28→Gln	++	-	++
Glu28→Lys	++	-	++
++ → binding equivalent to wildtype binding			
- → no binding			

7.4 ELISA analysis of antibody binding to residue 28 GST-MSP1₁₉ variants

ELISA was used to analyse the kinetics of antibody binding to GST-MSP1₁₉ variants over a range of antibody concentrations to look for subtle differences between the residue 28 variations that may not have been seen in the western blotting. The ELISA experiments were carried out using the ELISA conditions that were optimised for the analysis of the other amino acid variants described in chapter 3.

1 µg/ml wildtype and residue 28 GST-MSP1₁₉ variants were bound to 0.4 µg/ml goat-anti-GST antibody bound to the ELISA plate. The proteins were probed with doubling dilutions of B6, F5 or B10 antibody and 1/2000 dilution anti-mouse IgG HRP conjugate. The peroxidase was detected and absorbance was read at 490 nm. The ELISA data for F5 antibody (shown in figure 7.4) shows that both variations to residue 28 abolish F5 binding. The ELISA data for the B10 (shown in figure 7.5) and B6 (shown in figure 7.3) antibodies show that there is a very similar effect on antibody binding for the residue 28 variants with both changes reducing the steepness of the curve but the Glu28→Gln variation reduces the steepness of the curve to a lesser extent.

7.5 Expression and purification of Glu28→Gln his-MSP1₁₉ variant

In order to carry out NMR analysis of the Glu28→Gln MSP1₁₉ variant milligram quantities of his-tagged protein were required. The recodonised wildtype MSP1₁₉ gene in the pPIC9K vector was used as a template to produce a recodonised Glu28→Gln his-MSP1₁₉ gene using the XL-site directed mutagenesis kit. The recodonised Glu28→Gln MSP1₁₉ gene was cloned into pPIC9K (as described in section 2.4). Large scale expression of the Glu28→Gln his-MSP1₁₉ variant was carried out using the optimised protocol for expression derived for the other 5 his-MSP1₁₉ variants described in chapter 6.

7.6 ¹⁵N-HSQC NMR analysis of Glu28→Gln his-MSP1₁₉ variant

In order to identify differences in the protein structure of Glu28→Gln his-MSP1₁₉ variant compared to wildtype and Glu28→Lys his-MSP1₁₉ variant 2D ¹⁵N-HSQC NMR experiments were carried out. A ¹⁵N-HSQC spectrum was acquired at 25 °C for Glu28→Gln his-MSP1₁₉ variant as described in section 2.5.2 of materials and methods. The spectrum was overlayed on top of the wildtype spectrum using SPARKY software (Goddard & Kneller). The spectra were compared to identify NH peaks that had moved. Since a large number of peaks had moved in comparison to the wildtype spectrum (as shown in figure 7.6) the NH peaks that had moved could not be accurately assigned by comparing to the wildtype and assuming the peak corresponded to the closest assigned wildtype peak. The Glu28→Gln his-MSP1₁₉ variant spectrum was therefore assigned by acquiring ¹⁵N-HSQC-NOESY spectra at 25 °C and comparing the peak patterns with the ¹⁵N-HSQC-NOESY spectra for wildtype MSP1₁₉ and Glu28→Lys his-MSP1₁₉ variant (discussed in chapter 9) to confirm the identities of NH peaks. Figure 7.6 shows the spectrum for Glu28→Gln MSP1₁₉ variant in green overlayed on top of the wildtype spectrum shown in red (top panel). This clearly shows that the two spectra are quite different with a large number of NH peaks moving in both the ¹⁵N and ¹H dimensions. The peaks that have moved in the Glu28→Gln MSP1₁₉ variant spectrum can be seen in the histogram (bottom panel). This indicates that there are a number of clusters of

movement including between residues 5 and 14, 19 and 23, 27 and 31 and residue 95. The residues corresponding to NH peak movements of 0.05 ppm or more in the spectrum were mapped onto the 3D structure of wildtype *P. yoelii* MSP1₁₉ (the production of the 3D structure of wildtype *P. yoelii* MSP1₁₉ will be discussed in chapter 9) as shown in figure 7.7. This shows that the residues corresponding to the NH peaks that have moved in the Glu28→Gln MSP1₁₉ variant spectrum correspond to residues that are covering a large part of the first EGF domain and part of the second EGF domain and include residues that are not in direct contact with residue 28.

The ¹⁵N-HSQC spectrum for Glu28→Gln MSP1₁₉ variant was compared to the ¹⁵N-HSQC spectrum for Glu28→Lys MSP1₁₉ variant to see if there was less movement. Figure 7.8 shows the Glu28→Gln MSP1₁₉ variant spectrum in green overlayed on top of the Glu28→Lys MSP1₁₉ variant spectrum in blue and wildtype MSP1₁₉ in red. This shows that there are areas of the Glu28→Gln MSP1₁₉ variant spectrum where the peaks are in-between the wildtype peaks and the Glu28→Lys MSP1₁₉ variant peaks suggesting that some of the peaks have not moved as far in the Glu28→Gln MSP1₁₉ variant spectrum. The histogram in figure 7.9 shows the combined chemical shift difference for the ¹H and ¹⁵N dimension for the Glu28→Gln MSP1₁₉ variant spectrum in green and Glu28→Lys MSP1₁₉ variant spectrum in blue. This shows that the overall pattern of the histogram is the same suggesting that the same residues have moved. The histogram in figure 7.10 compares the difference between the combined chemical shift difference for the Glu28→Gln and Glu28→Lys MSP1₁₉ variants. In this histogram, the combined chemical shift difference for Glu28→Gln MSP1₁₉ variant has been taken away from the Glu28→Lys MSP1₁₉ variant combined chemical shift difference. The bars with positive values therefore indicate NHs that have moved more in the Glu28→Lys MSP1₁₉ variant and the bars with negative values indicate NHs that have moved more in the Glu28→Gln MSP1₁₉ variant. This histogram shows that overall the chemical shift differences are larger for the Glu28→Lys MSP1₁₉ variant and show the most difference for residues 9, 10, 13, 21, 22, and 29 to 31. This histogram also shows that for residues 11, 12 and 28 there is a larger chemical shift difference for Glu28→Gln MSP1₁₉ variant.

Figure 7.2: Western blotting analysis of antibody binding to residue 28 GST-MSP1₁₉ variants.

500 ng of the wildtype GST-MSP1₁₉ and residue 28 GST-MSP1₁₉ variants were run on NuPAGE gels under reducing conditions (anti-GST antibody western blots) or non-reducing conditions (B6, F5 and B10 antibody western blots) and transferred to nitrocellulose. GST and BSA were used as controls. The anti-GST western blot (top panel) was probed with 1/5000 dilution anti-GST HRP conjugate. The B6, F5 and B10 antibody western blots were first probed with B6 (2 µg/ml), F5 (10 µg/ml) or B10 (2 µg/ml) followed by 1/2000 dilution goat anti-mouse IgG HRP conjugate. The bands at 37 kDa are the GST-MSP1₁₉ proteins.

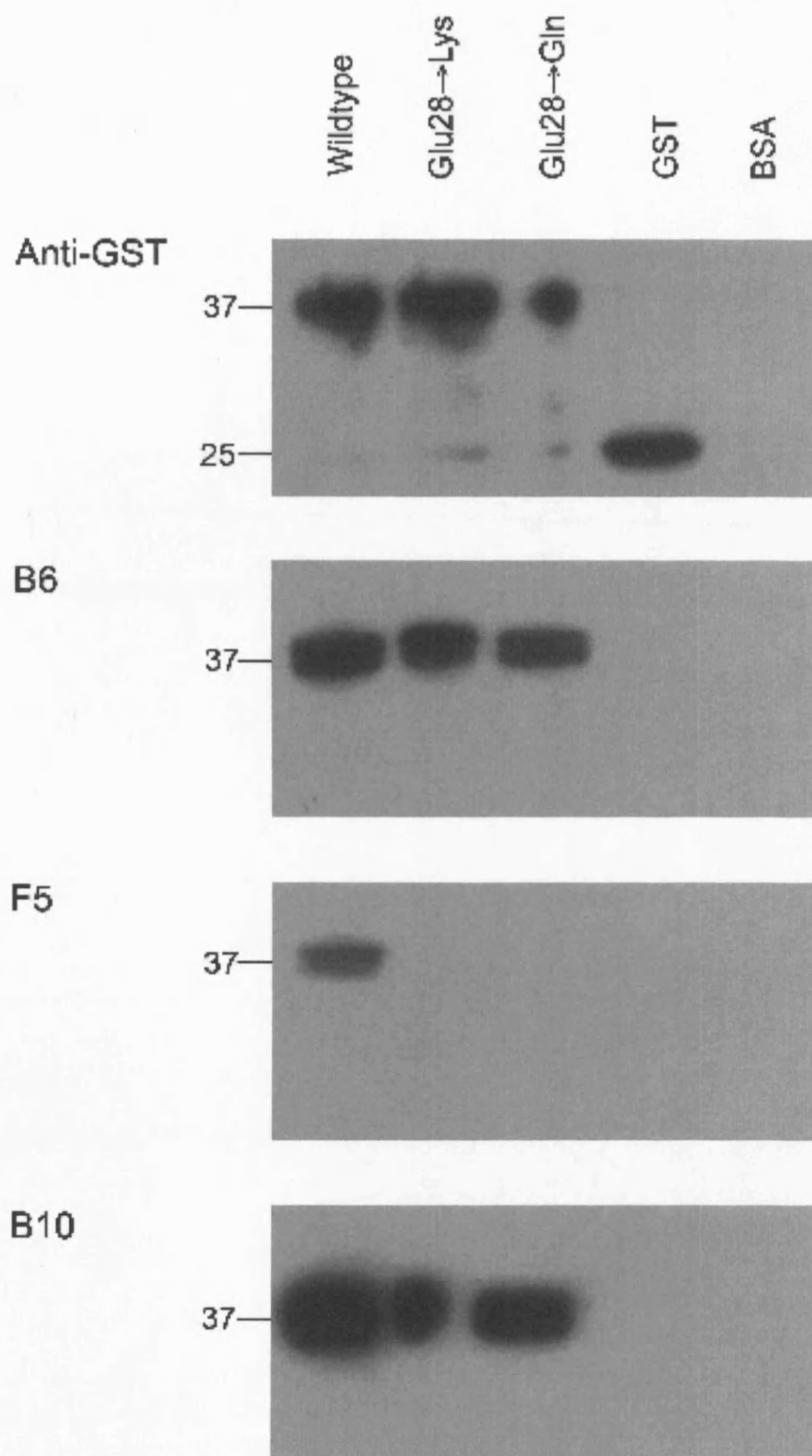


Figure 7.3: ELISA of B6 antibody binding to residue 28 GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype GST-MSP1₁₉, residue 28 GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B6 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in blue and Glu28→Gln GST-MSP1₁₉ variant is shown in green.

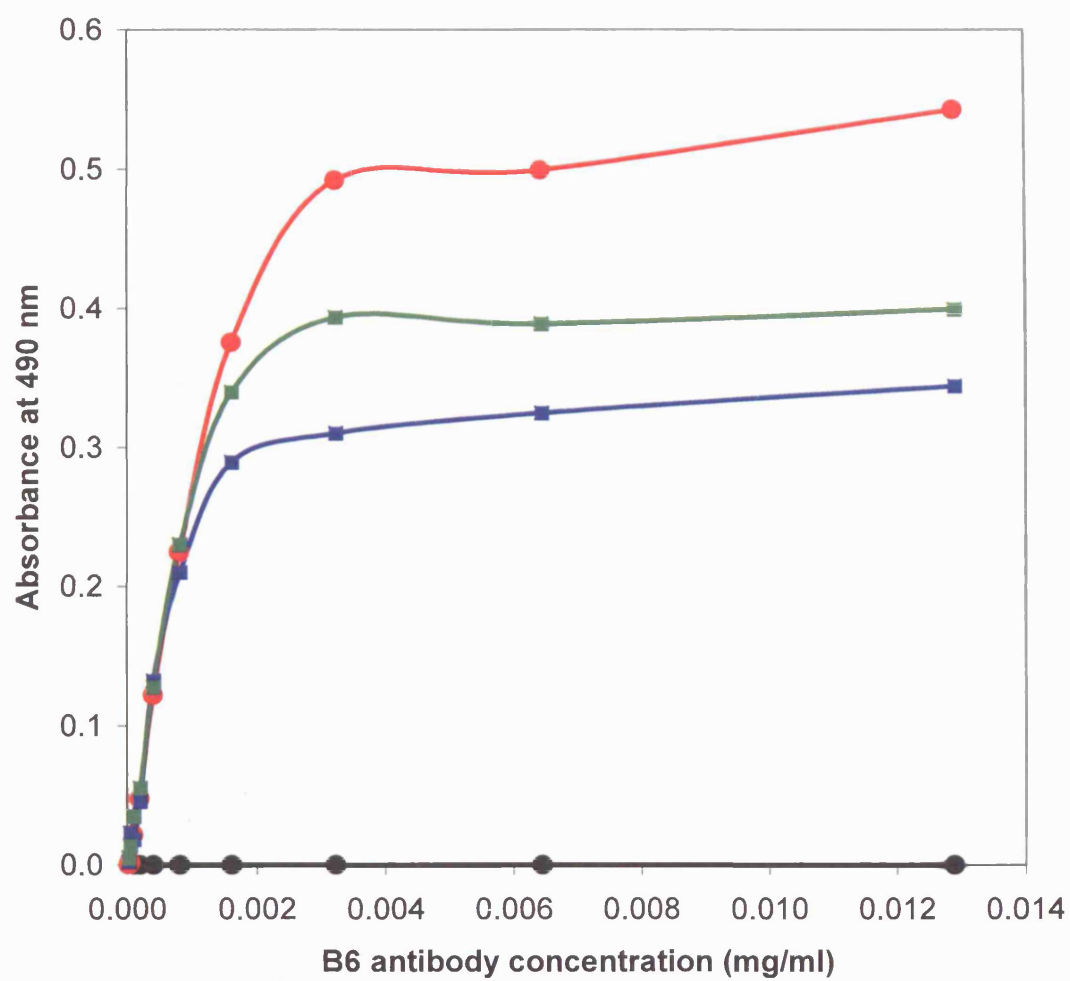


Figure 7.4: ELISA of F5 antibody binding to residue 28 GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype GST-MSP1₁₉, residue 28 GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of F5 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in blue and Glu28→Gln GST-MSP1₁₉ variant is shown in green.

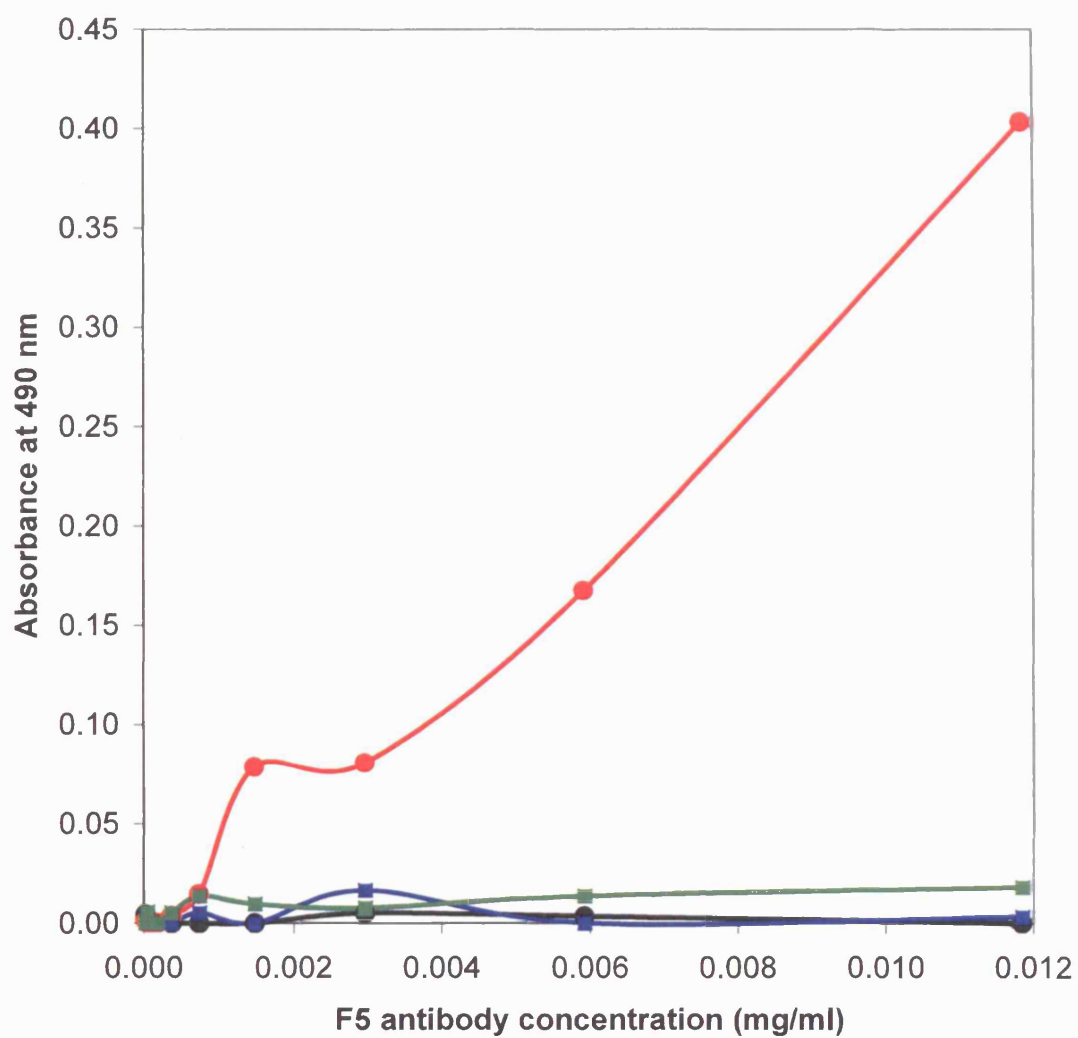


Figure 7.5: ELISA of B10 antibody binding to residue 28 GST-MSP1₁₉ variants.

0.4 µg/ml goat anti-GST antibody was used to coat the ELISA plate as the capture antibody. 1 µg/ml of wildtype GST-MSP1₁₉, residue 28 GST-MSP1₁₉ variants, GST or PBS was bound to the capture antibody. This was probed with doubling dilutions of B10 antibody and 1/1000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the variant proteins and GST less PBS control are shown on the graph. GST is shown in black, wildtype GST-MSP1₁₉ is shown in red, Glu28→Lys GST-MSP1₁₉ variant is shown in blue and Glu28→Gln GST-MSP1₁₉ variant is shown in green.

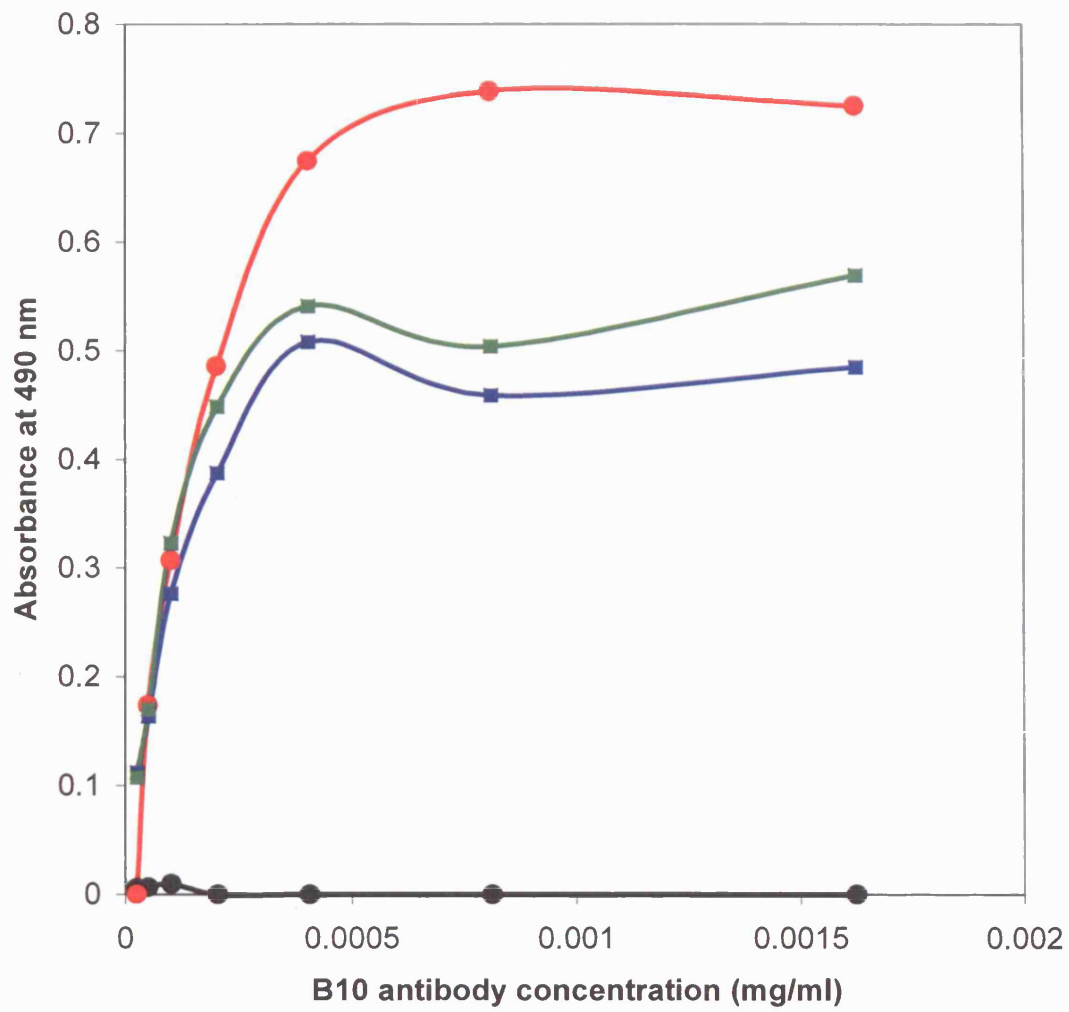


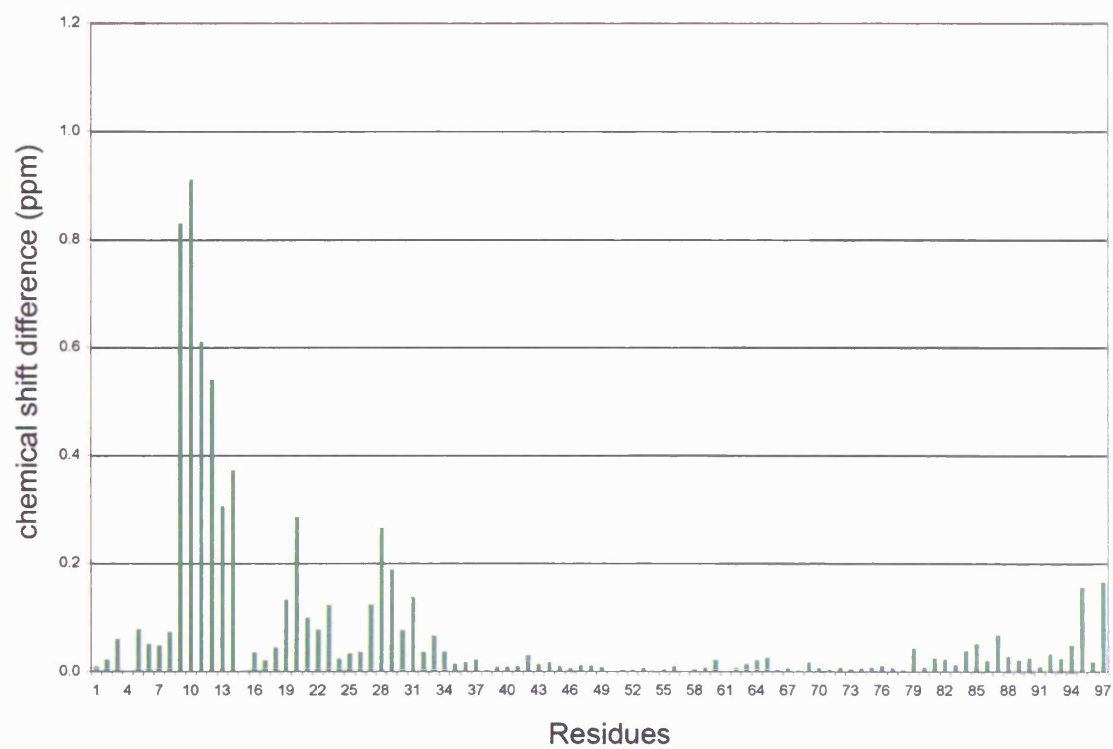
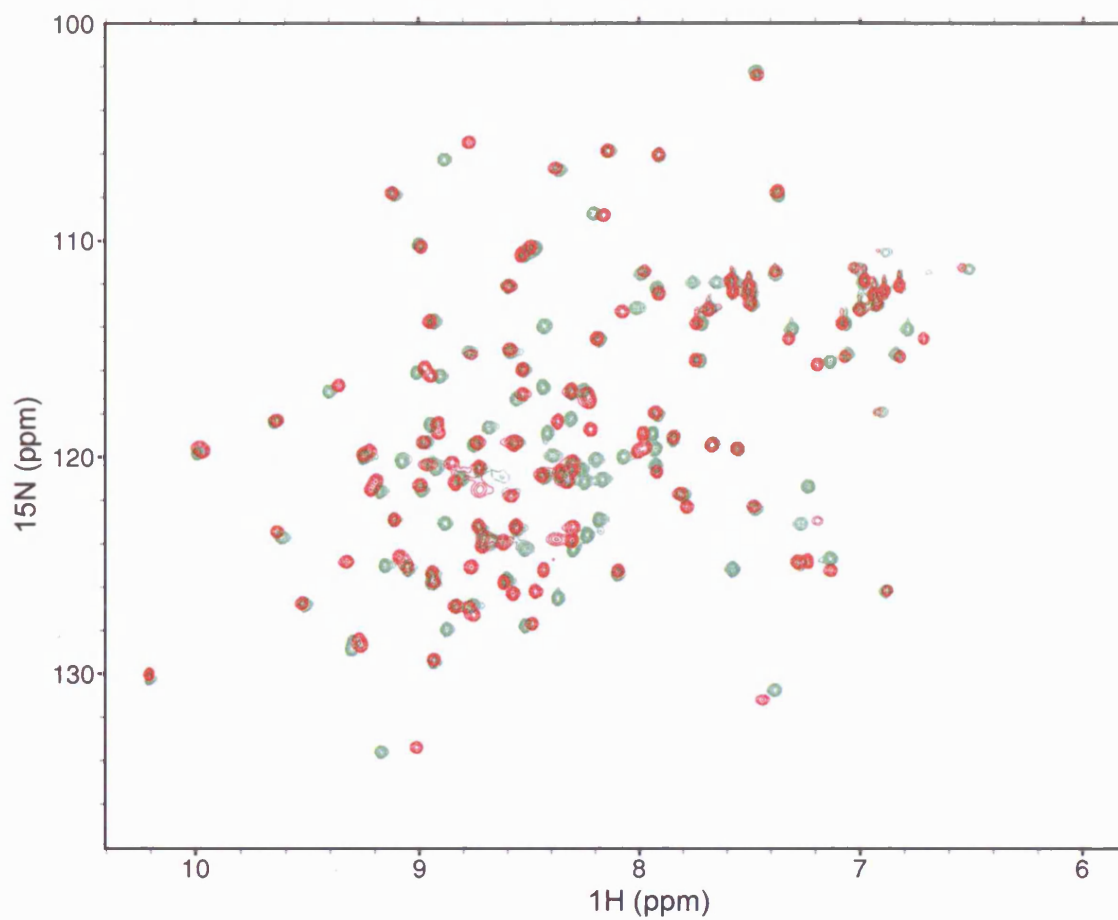
Figure 7.6: Comparison of Glu28→Gln MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

Top: The ¹⁵N-HSQC spectrum of Glu28→Gln MSP1₁₉ variant is shown in green overlayed on top of the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software (Goddard & Kneller). The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

Bottom: Histogram showing the combined ¹⁵N and ¹H chemical shift differences for the Glu28→Gln MSP1₁₉ variant ¹⁵N-HSQC spectrum compared to the wildtype spectrum.

The following equation was used to combine the chemical shifts:

Absolute ((variant ¹⁵N chemical shift – wildtype ¹⁵N chemical shift) ÷ 5) + Absolute (variant ¹H chemical shift – wildtype ¹H chemical shift) ÷ 2



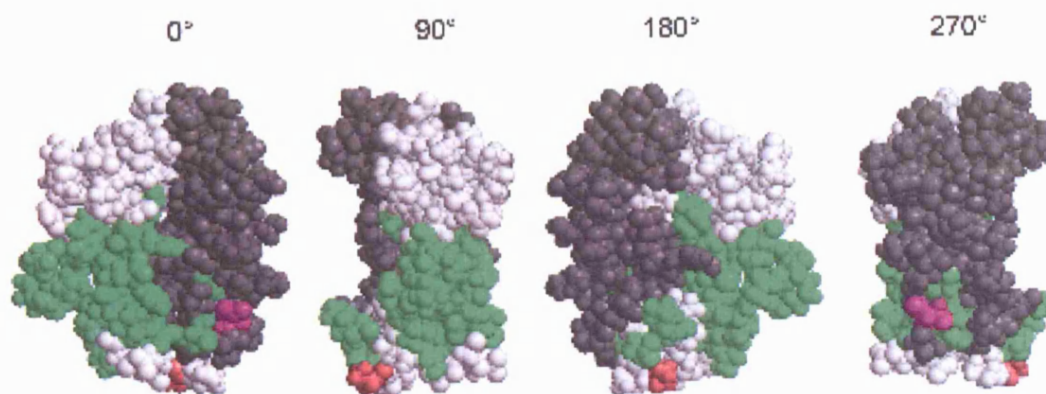


Figure 7.7: Mapping the NH peaks that have moved in the Glu28→Gln MSP1₁₉ variant ¹⁵N-HSQC spectra onto the 3D structure of *P. yoelii* MSP1₁₉.

The residues corresponding to the NH peaks that have moved more than 0.05 ppm for the combined chemical shift difference in the ¹⁵N-HSQC spectra for the variants compared to the wildtype MSP1₁₉ have been mapped onto the best energy NMR structure for the wildtype *P. yoelii* MSP1₁₉. The calculation of this structure is discussed in chapter 9. The residue 28 is shown in red. The residues corresponding to the NH peaks that have moved more than 0.05 ppm in the Glu28→Gln MSP1₁₉ variant ¹⁵N-HSQC spectra are shown in green. The first EGF domain is shown in white and the second EGF domain is shown in grey. The C-terminal residue is shown in bright pink and the N-terminal residue is shown in light pink.

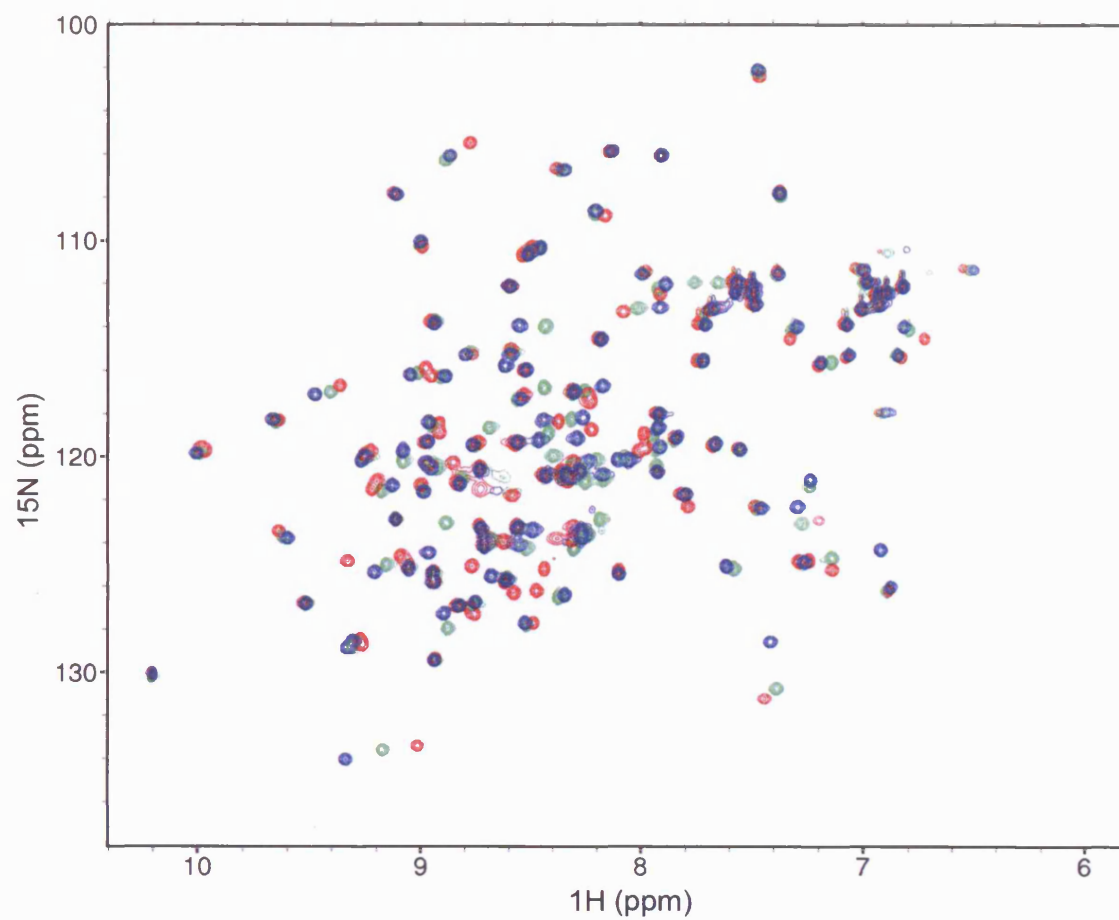


Figure 7.8: Comparison of Glu28→Gln MSP1₁₉ variant, Glu28→Lys MSP1₁₉ variant and wildtype MSP1₁₉ ¹⁵N-HSQC NMR spectra.

The ¹⁵N-HSQC spectrum of Glu28→Gln MSP1₁₉ variant is shown in green overlayed on top of the Glu28→Gln MSP1₁₉ variant (shown in blue) and the wildtype MSP1₁₉ spectrum (shown in red) using SPARKY software. The spectra were acquired at 600 MHz at 25 °C in 25 mM potassium phosphate buffer, 50 mM KCl, pH 6.5.

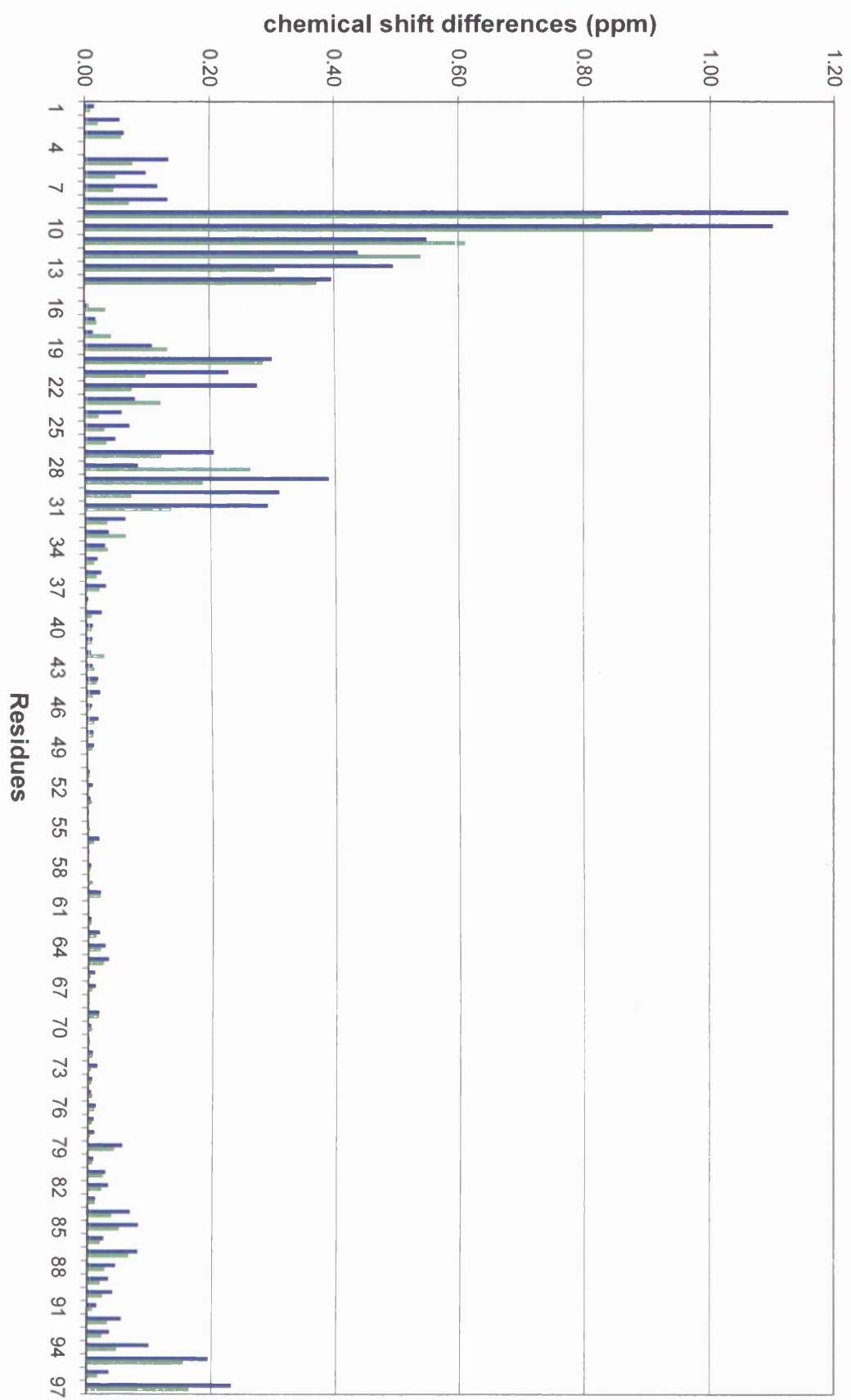


Figure 7.9: Comparison of the combined ^{15}N and ^1H chemical shift differences of Glu28→Gln and Glu28→Lys MSP1₁₉ variants compared to wildtype MSP1₁₉.

Histogram showing the combined ^{15}N and ^1H chemical shift differences for the Glu28→Gln and Glu28→Lys MSP1₁₉ variants ^{15}N -HSQC spectra compared to the wildtype spectrum. The Glu28→Gln MSP1₁₉ variant is shown in green and the Glu28→Lys MSP1₁₉ variant is shown in blue.

The following equation was used to combine the chemical shifts:

Absolute $((\text{variant } ^{15}\text{N chemical shift} - \text{wildtype } ^{15}\text{N chemical shift}) \div 5) + \text{Absolute} (\text{variant } ^1\text{H chemical shift} - \text{wildtype } ^1\text{H chemical shift}) \div 2$

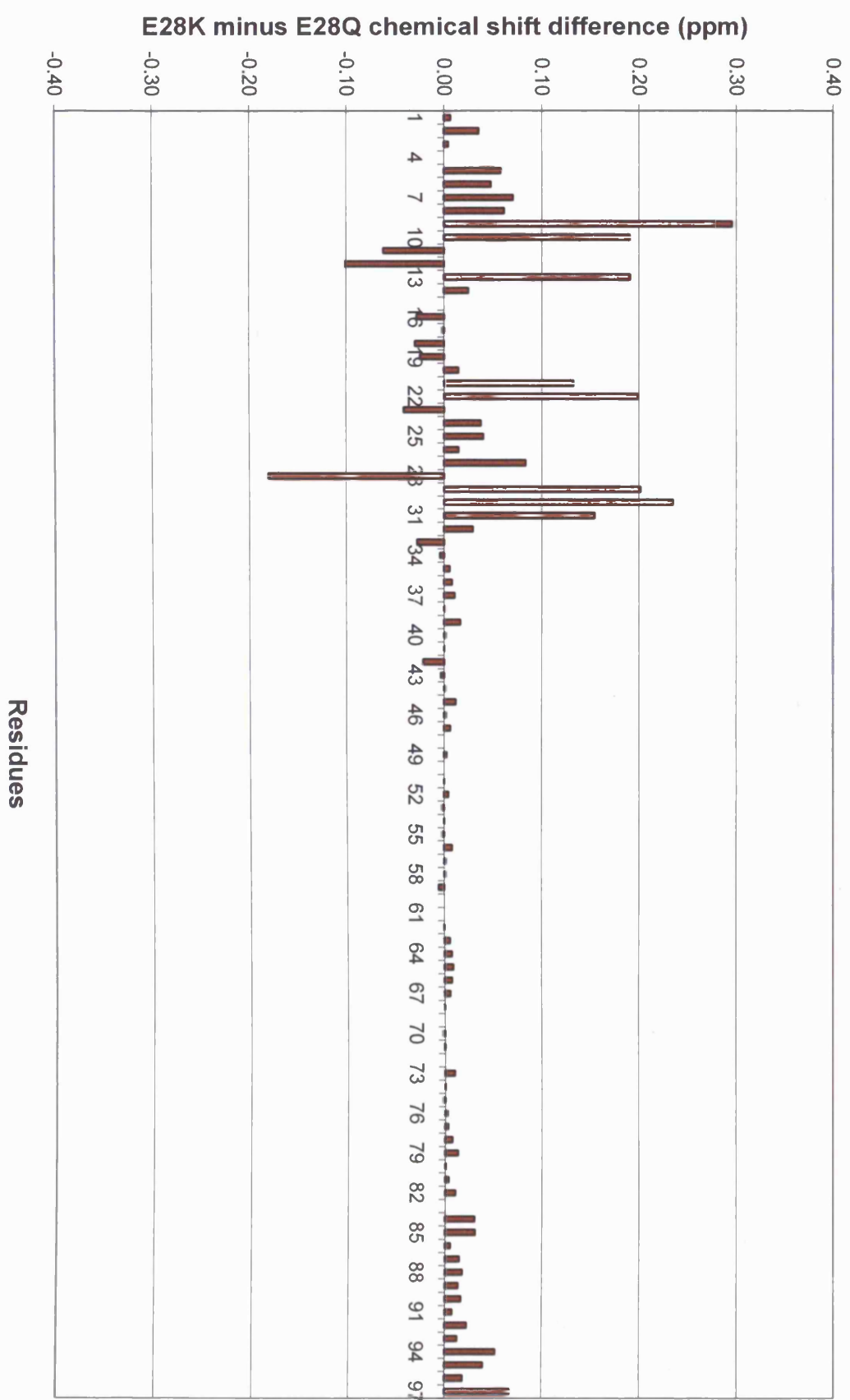


Figure 7.10: Differences between the combined ^{15}N and ^1H chemical shift differences of Glu28→Gln and Glu28→Lys MSP1₁₉ variants compared to wildtype MSP1₁₉.

Histogram showing the difference between the combined ^{15}N and ^1H chemical shift differences for the Glu28→Gln and Glu28→Lys MSP1₁₉ variants ^{15}N -HSQC spectra compared to the wildtype spectrum. The positive values indicate where the chemical shift difference is larger for the Glu28→Lys MSP1₁₉ variant ^{15}N -HSQC spectrum. The negative values indicate where the combined chemical shift difference is larger for Glu28→Gln MSP1₁₉ variant ^{15}N -HSQC spectrum.

The following equation was used to combine the chemical shifts:

$$\text{Absolute } ((\text{variant } ^{15}\text{N chemical shift} - \text{wildtype } ^{15}\text{N chemical shift}) \div 5) + \text{Absolute } (\text{variant } ^1\text{H chemical shift} - \text{wildtype } ^1\text{H chemical shift}) \div 2$$

The following equation was used to calculate the difference between the combined chemical shifts:

$$\text{Glu28} \rightarrow \text{Lys MSP1}_{19} \text{ variant combined chemical shift} - \text{Glu28} \rightarrow \text{Gln MSP1}_{19} \text{ variant combined chemical shift.}$$

7.7 Discussion

The data presented in this chapter have shown that overall the Glu28→Gln MSP1₁₉ variant affects antibody binding in the same way as the Glu28→Lys MSP1₁₉ variant. This suggests that even small changes with just a loss of charge can affect antibody binding.

The ¹⁵N-HSQC NMR spectrum for Glu28→Gln was visually very different to the wildtype with most of the NH peaks in the first domain moving and some by a large amount. This suggests that there is a lot of structural perturbation as a result of the Glu28→Gln variation. This means that I cannot be confident that the differences in antibody binding were a direct result of the change to residue 28 and not the structural changes as a result of the variation. The comparison of ¹⁵N-HSQC NMR spectrum of Glu28→Gln MSP1₁₉ variant to the Glu28→Lys MSP1₁₉ variant showed a very similar pattern of peak movement. The distance of peak movement was generally less for the Glu28→Gln MSP1₁₉ variant suggesting there is less structural perturbation as a result of Glu28→Gln variation. Five of the residues that have moved significantly more in the Glu28→Lys MSP1₁₉ variant are charged residues. The charged residues may have moved more in the Glu28→Lys MSP1₁₉ variant because the positively charged residues may be repelled by the lysine in the variant, whereas they could have been attracted in the wildtype. The negatively charged residues may have been repelled by glutamic acid in the wildtype but now attracted by lysine in the variant. The Glu28→Gln MSP1₁₉ variant may therefore not have had such a significant affect on the charged residues when compared to Glu28→Lys MSP1₁₉ variant because the change to an uncharged residue would not repel or attract the charged residues. This could mean that for the Glu28→Gln MSP1₁₉ variant any charged residues that were being attracted by the glutamic acid may not be attracted as closely but would not be pushed away which could lead to less change in the chemical environment of the residues and therefore less movement in the spectrum.

The larger change in peak position for residue 12 in the Glu28→Gln MSP1₁₉ is difficult to explain. Residue 12 is a positively charged residue and I would have predicted a

larger change in the Glu28→Lys MSP1₁₉ variant because the lysine may have repelled this residue. The larger change in peak position for residue 12 in Glu28→Gln MSP1₁₉ variant may be because the residue has moved to a chemical environment that is more different than the environment it has moved to in the Glu28→Lys MSP1₁₉ variant resulting in a big shift in the spectrum. The larger change in peak position for residue 28 in the Glu28→Gln MSP1₁₉ variant spectrum may be because the positions of a glutamic acid residue and a glutamine residue in a ¹⁵N-HSQC spectrum could be further apart than the positions of a glutamic acid and a lysine residue in a ¹⁵N-HSQC spectrum.

The differences in the positions of the peaks in the ¹⁵N-HSQC spectra can only be an indicator of the level of structural perturbation and cannot identify the exact changes to the structure. I can therefore conclude that it is likely that the Glu28→Gln MSP1₁₉ variant has caused less changes to the structure as the level of structural perturbation is less but cannot confirm the exact changes.

Chapter 8: Further immunisation studies 2 – do the residue 28 variants affect protection differently?

8.1 Introduction

The immunisation studies in chapter 5 showed that immunisation with Glu28→Lys and double Lys16→Glu/Glu28→Lys MSP1₁₉ variants did not protect against subsequent challenge infection with *P. yoelii* YM. Wildtype MSP1₁₉ and the residue 12, 16 and 17 variants were able to protect against challenge infection. The 2D ¹⁵N-HSQC NMR studies in chapter 6 showed that the Glu28→Lys variation caused a large amount of structural perturbation to the protein. This meant that I could not be confident that the affect on *in vivo* protection observed for the Glu28→Lys MSP1₁₉ variant was a direct result of the variation to residue 28. A Glu28→Gln MSP1₁₉ variant was created and antibody binding was analysed (as described in chapter 7). This showed that the Glu28→Gln MSP1₁₉ variant had a very similar effect on antibody binding *in vitro* as the Glu28→Lys MSP1₁₉ variant. 2D ¹⁵N-HSQC NMR studies on the Glu28→Gln MSP1₁₉ variant suggested that this change caused less structural perturbation than the Glu28→Lys MSP1₁₉ variant.

In this chapter, I will discuss immunisation studies to look at the affect on protection of the residue 28 variants. The studies will compare the affect on *in vivo* protection of the Glu28→Lys, Glu28→Gln and the double Lys16→Glu/Glu28→Lys MSP1₁₉ variants. There could be a number of outcomes of the immunisation studies: the Glu28→Gln MSP1₁₉ variant could have the same affect on protection as the Glu28→Lys and the double Lys16→Glu/Glu28→Lys MSP1₁₉ variants; the Glu28→Gln MSP1₁₉ variant could have less of an affect on protection than the Glu28→Lys and double Lys16→Glu/Glu28→Lys MSP1₁₉ variants or the Glu28→Gln MSP1₁₉ could have no affect on protection.

ELISA experiments will also be discussed in this chapter to compare the antibody titres from immunisation with wildtype and residue 28 MSP1₁₉ variants versus wildtype

MSP1₁₉ protein and residue 28 variant proteins. The ELISA studies will examine if there is a difference in the fine specificity of the antibody responses to the wildtype and residue 28 MSP1₁₉ variants that could explain any differences in protection.

8.2 Immunisation studies with residue 28 MSP1₁₉ variants

The immunisation studies were carried out as described in materials and methods (section 2.3). Immunisation studies were carried out with the five single amino acid variants (Arg12→Leu, Lys16→Glu, Asn17→His, Glu28→Lys, Glu28→Gln) and one double amino acid GST-MSP1₁₉ variant (Lys16→Glu/Glu28→Lys). Wildtype GST-MSP1₁₉ was used as a positive control and purified GST was used as a negative control. Following immunisation with the GST-MSP1₁₉ variants, the mice were challenged with *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily on Giemsa stained blood films. The percentage parasitaemia was determined to compare between the variants. The overall results of the immunisation studies (data not shown) confirmed the results obtained in the first immunisation studies in chapter 5 without the Glu28→Gln MSP1₁₉ variant. In this chapter, I will therefore only focus on the results of the immunisation studies for the residue 28 MSP1₁₉ variants (and wildtype MSP1₁₉ and GST as the controls).

The graphs in panel A of figures 8.1 to 8.5 show the parasitaemia for each of the mice in the groups. Figure 8.1, panel A, shows the parasitaemia for the mice immunised with wildtype GST-MSP1₁₉. This shows that five of the mice immunised with wildtype GST-MSP1₁₉ were able to clear the parasites. Two of the mice had very low parasitaemia and three had higher parasitaemia. One of the mice was unable to clear the parasites and was killed by a schedule one method on day 17. Figure 8.2 shows the parasitaemia for the mice immunised with GST as negative controls. This shows a rapid increase in parasitaemia up to day 6 and all the mice were killed by a schedule one method on day 7. Figure 8.3, panel A, shows the parasitaemia for the mice immunised with Glu28→Lys MSP1₁₉ variant. This shows that four of the mice had a rapid increase in parasitaemia up to day 6 or 7 and were killed by a schedule one method on day 6 or 7. One mouse had a lower parasitaemia but was killed by a schedule one method on day 13 due to the severity of malaria symptoms. One mouse was able to clear the parasites.

Figure 8.4, Panel A, shows the parasitaemia for the mice immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant. This shows that five of the mice had a rapid increase in parasitaemia and were killed by a schedule one method. One mouse had high parasitaemia but was able to clear the parasites. Figure 8.5, panel A, shows the parasitaemia for the mice immunised with Glu28→Gln MSP1₁₉ variant. This shows that four of the mice had a rapid increase in parasitaemia and were killed by a schedule one method. The other two mice had very low parasitaemia and were able to clear the parasites.

Figure 8.6, panel A, shows the average parasitaemia of the six mice in the groups. The average parasitaemia is less clear than the individual parasitaemia counts seen in figures 8.1 to 8.5 because in each of the groups the parasitaemia for at least one mouse did not follow the parasitaemia for the rest of the group. The overall results suggest that the parasitaemia is higher in the early days of the infection for the mice immunised with the residue 28 variants than the mice immunised with the wildtype variants.

8.3 ELISA analysis of antibody titres following immunisation with residue 28 MSP1₁₉ variants

The antibody titres following immunisation with the GST-MSP1₁₉ variants were compared to those of the wildtype GST-MSP1₁₉ to determine whether any differences in protection from parasite challenge could be explained by differences in the level of antibody response to the GST-MSP1₁₉ variants. The ELISA experiments were carried out as described in materials and methods (section 2.3.1). In the ELISA experiments his-tagged wildtype MSP1₁₉ was used to analyse the antibody levels instead of GST-MSP1₁₉. This was to avoid problems associated with the production of antibodies to the GST portion of GST-MSP1₁₉ which could saturate the ELISA signal making it difficult to see small differences in antibody titre to the MSP1₁₉ portion. 1 µg/ml wildtype his-MSP1₁₉ was bound to the ELISA plate (the production of his-MSP1₁₉ is described in section 2.4). The proteins were probed with doubling dilutions of pooled serum samples from the six mice in the groups and 1/2000 dilution anti-mouse IgG HRP conjugate. The peroxidase was detected and absorbance was read at 490 nm. The ELISA results are shown in figure 8.6, panel B. The ELISA results show that all the mice immunised

with all of the GST- MSP1₁₉ variants had produced antibodies to MSP1₁₉. The binding curve for the mice immunised with GST shows that the ELISA system used is specific for antibody binding to MSP1₁₉. The ELISA results show that there are no significant differences between the pooled antibody titres for the mice immunised with the residue 28 variant MSP1₁₉ proteins and wildtype MSP1₁₉. The average absorbance difference between the residue 28 variants and the wildtype MSP1₁₉ was 0.13.

ELISA experiments were also carried out with the serum from each individual mouse in the group using the same experimental conditions as were used for testing the pooled sera to examine whether the results of the pooled sera were an accurate reflection of the results obtained with the serum from the individual mice. The ELISA results for serum from the individual mice may also explain why some of the mice in a group had high parasitaemia while others had low parasitaemia. The ELISA results for the serum from individual mice immunised with wildtype MSP1₁₉ are shown in figure 8.1, panel B. The results show that all of the mice immunised with wildtype MSP1₁₉ have produced antibodies to wildtype MSP1₁₉ and that all of the antibodies titres are very similar. The ELISA results for the serum from individual mice immunised with Glu28→Lys MSP1₁₉ variant are shown in figure 8.3, panel B. The ELISA results show that there is a significant difference between the antibody binding curves for the individual mice immunised with Glu28→Lys MSP1₁₉. The ELISA results for the serum from individual mice immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant are shown in figure 8.4, panel B. The ELISA results show that there are differences in the antibody titres with a spread of antibody titres from mouse two and four having lower antibody titres and mouse one having the highest antibody titre to wildtype MSP1₁₉. The ELISA results for the serum from individual mice immunised with Glu28→Gln MSP1₁₉ variant are shown in figure 8.5, panel B. The ELISA results show that there are differences in the antibody titres for the individual mice with a narrow spread of antibody titres from mouse four having the highest antibody titre and mouse one and two having the lowest antibody titres to wildtype MSP1₁₉.

8.4 ELISA analysis of serum antibodies against wildtype MSP1₁₉ and residue 28 MSP1₁₉ variants

ELISA experiments were carried out binding Glu28→Lys, Glu28→Gln, and double Lys16→Glu/Glu28→Lys MSP1₁₉ variants to the ELISA plate and probing with doubling dilutions of the serum from the individual mice immunised with Glu28→Lys, Glu28→Gln, and double Lys16→Glu/Glu28→Lys MSP1₁₉ variants. The experiments were carried out to compare the antibody titres to the residue 28 variants and the wildtype protein. This could confirm if there were lots of antibodies being made to the area that had been changed in the residue 28 variants. If a large proportion of the antibody response had been to the area of the protein that had been altered in the residue 28 variants, I would expect a difference in the antibody titre for the serum antibodies binding to the residue 28 variants compared to the wildtype MSP1₁₉.

Figure 8.7, panel A shows the antibody binding curves for the serum from the individual mice immunised with Glu28→Lys MSP1₁₉ binding to his-tagged Glu28→Lys MSP1₁₉. This shows that there is a spread of antibody titres for the serum of the individual mice that follows the same pattern as the antibody titres to wildtype MSP1₁₉. Figure 8.7, panel B shows the antibody binding curves for the sera from the individual mice immunised with Glu28→Lys MSP1₁₉ binding to his-tagged Glu28→Lys MSP1₁₉ (outline shapes) and binding to his-tagged wildtype MSP1₁₉ (filled in shapes). This shows the antibody titres to the Glu28→Lys MSP1₁₉ and wildtype MSP1₁₉ proteins are different. For all the mice apart from mouse four the antibody titres to the wildtype MSP1₁₉ protein are lower than to the Glu28→Lys MSP1₁₉ protein. For mouse four the antibody titres are very similar to both proteins. The biggest difference between antibody binding to the two proteins is for mouse three. Figure 8.8, panel A shows the antibody binding curves for the serum from the individual mice immunised with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant binding to his-tagged double Lys16→Glu/Glu28→Lys MSP1₁₉ variant. This shows that there is a spread of antibody titres for the serum of the individual mice that is less spread out than the titres seen to wildtype MSP1₁₉ but follows the same pattern. Figure 8.7, panel B shows the antibody binding curves for the serum from the individual mice immunised with double

Lys16→Glu/Glu28→Lys MSP1₁₉ variant binding to his-tagged double

Lys16→Glu/Glu28→Lys MSP1₁₉ variant (outline shapes) and binding to his-tagged wildtype MSP1₁₉ (filled in shapes). This shows the antibody titres to the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant are higher than the antibody titres to the wildtype MSP1₁₉. Figure 8.9, panel A shows the antibody binding curves for the serum from the individual mice immunised with Glu28→Gln MSP1₁₉ variant binding to his-tagged Glu28→Gln MSP1₁₉ variant. This shows that there is a spread of antibody titres for the serum of the individual mice to Glu28→Gln MSP1₁₉ variant that is more spread out than the titres seen to wildtype MSP1₁₉. Figure 8.9, panel B shows the antibody binding curves for the sera from the individual mice immunised with Glu28→Gln MSP1₁₉ variant binding to Glu28→Gln MSP1₁₉ variant (outline shapes) and binding to his-tagged wildtype MSP1₁₉ (filled in shapes). This shows the antibody titres to the Glu28→Gln MSP1₁₉ variant are higher for mouse three to six than the antibody titres to the wildtype MSP1₁₉. The antibody binding titres to Glu28→Gln MSP1₁₉ variant for mouse one and two are very similar to the antibody titres to wildtype MSP1₁₉.

In order to look at the cross-reactivity of the sera from the mice immunised with wildtype MSP1₁₉ with the residue 28 proteins, ELISA experiments were carried out binding Glu28→Lys, Glu28→Gln and double Lys16→Glu/Glu28→Lys MSP1₁₉ variants to the ELISA plate and looking at binding of doubling dilutions of the pooled sera from the mice immunised with wildtype MSP1₁₉. The ELISA experiment would confirm if there were antibodies made to areas of the residue 28 variant proteins that were not altered by the variations. The results are shown in figure 8.10. The antibody binding curves for binding to wildtype MSP1₁₉ is shown in red, binding to Glu28→Lys MSP1₁₉ variant is shown in blue, binding to the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in orange and binding to Glu28→Gln MSP1₁₉ variant is shown in green. The ELISA shows that the antibody binding curves for the wildtype pooled sera to the wildtype MSP1₁₉ and the residue 28 variants are very similar. The antibody titre for the wildtype pooled sera to Glu28→Lys is slightly less than to the wildtype and other residue 28 variants.

Figure 8.1: Course of *P. yoelii* infection in mice immunised with wildtype MSP1₁₉ and antibody binding curves.

A: Six BALB/c mice were immunised with 10 µg of wildtype MSP1₁₉ in FCA followed by two injections with 40 µg of wildtype MSP1₁₉ in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with wildtype MSP1₁₉ and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

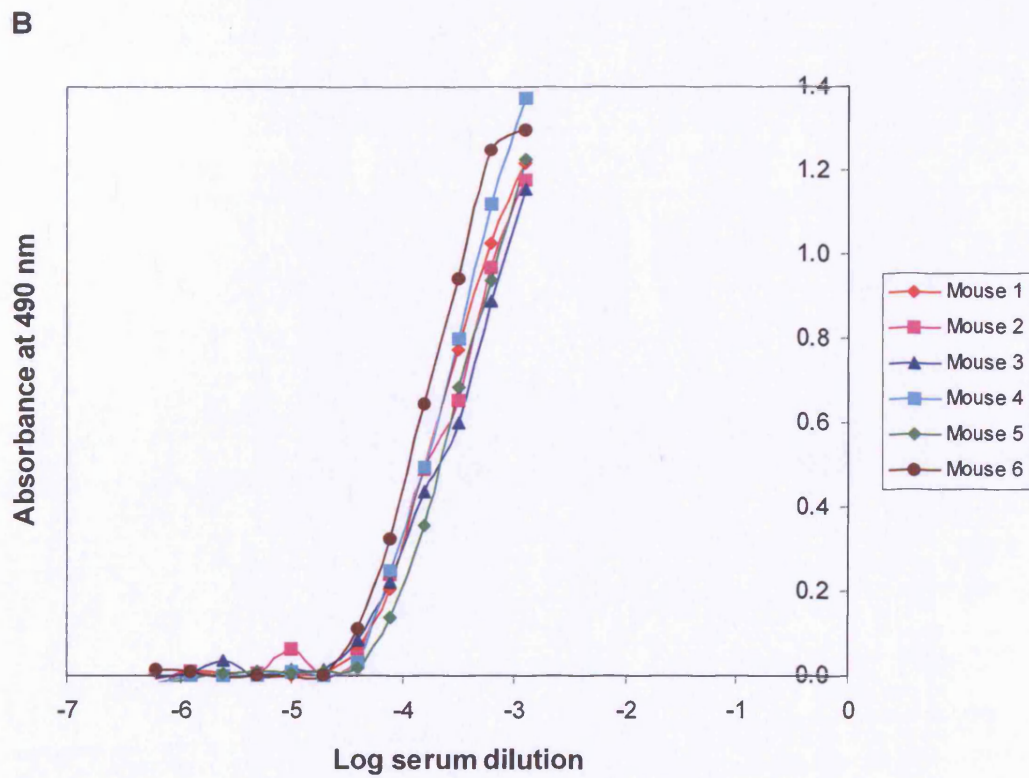
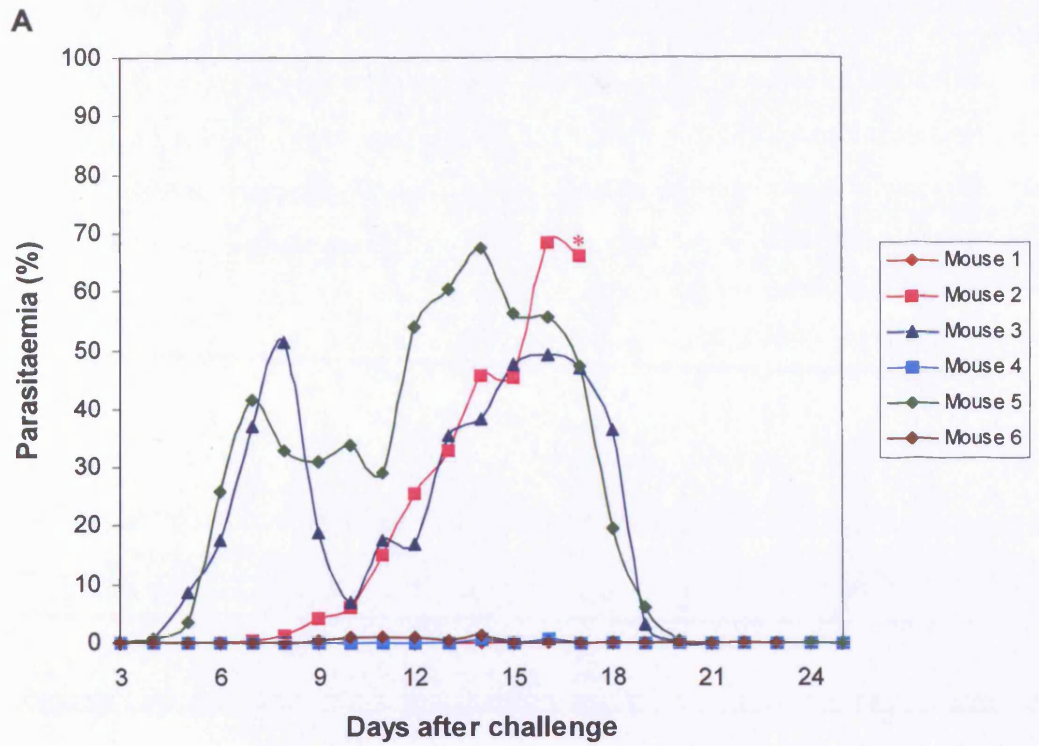


Figure 8.2: Course of *P. yoelii* infection in mice immunised with GST.

A: Six BALB/c mice were immunised with 10 µg of GST in FCA followed by two injections with 40 µg of GST in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

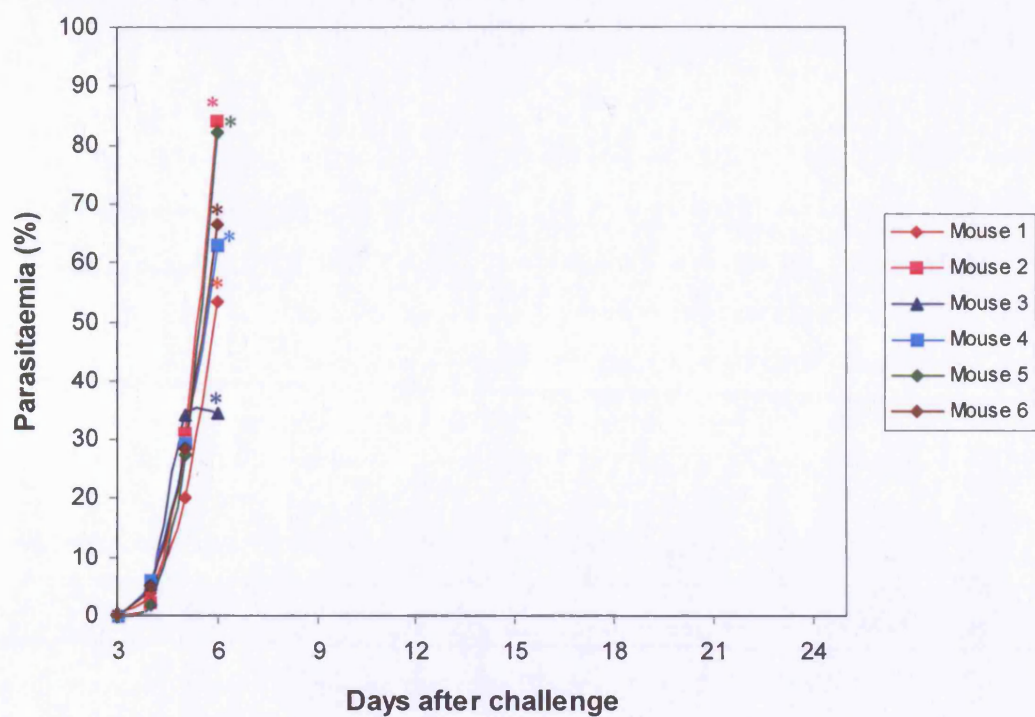


Figure 8.3: Course of *P. yoelii* infection in mice immunised with Glu28→Lys MSP1₁₉ variant and antibody binding curves.

A: Six BALB/c mice were immunised with 10 µg of Glu28→Lys MSP1₁₉ variant in FCA followed by two injections with 40 µg of Glu28→Lys MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

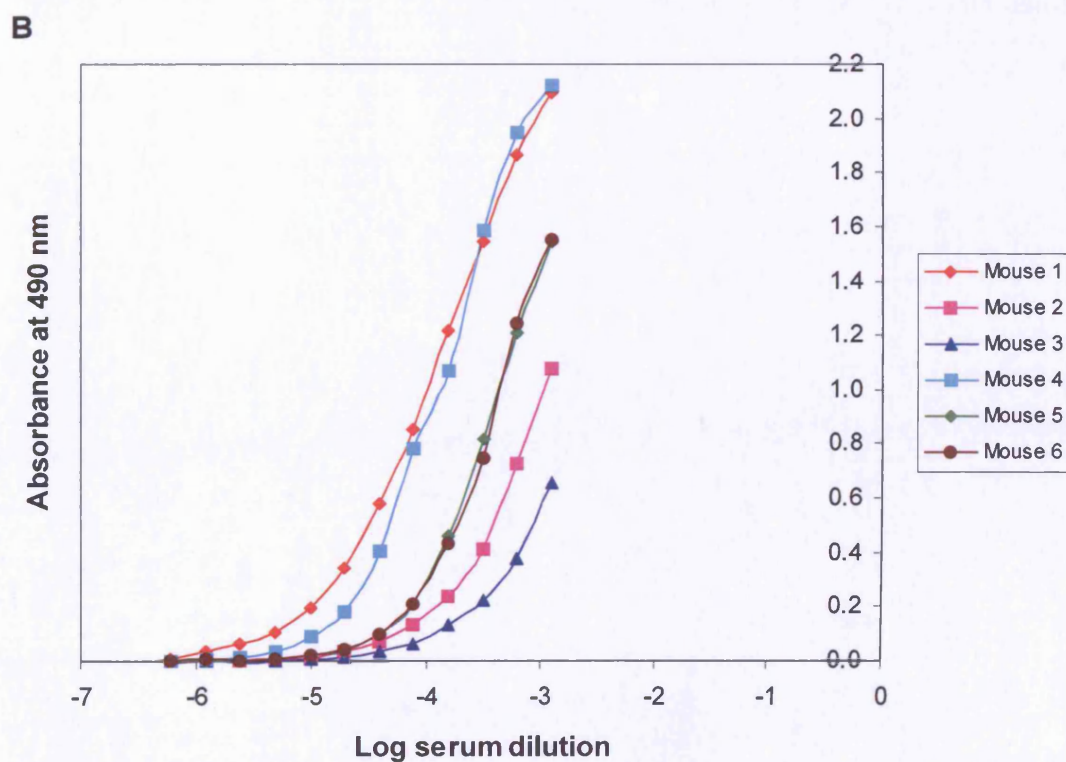
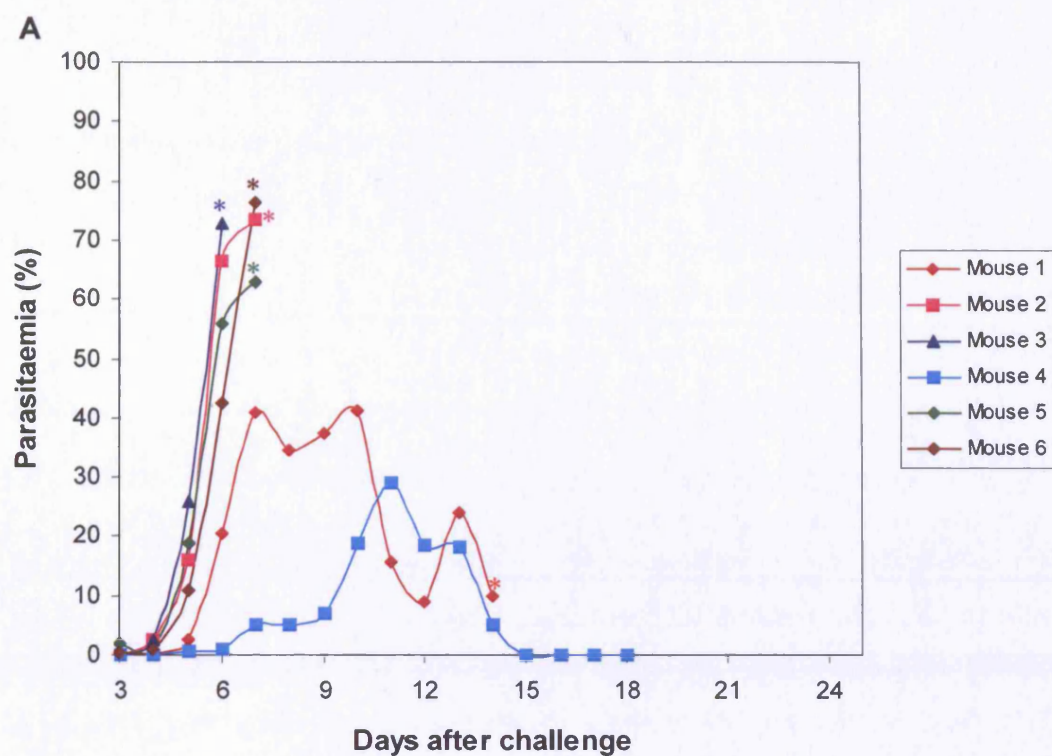


Figure 8.4: Course of *P. yoelii* infection in mice immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant and antibody binding curves.

A: Six BALB/c mice were immunised with 10 µg of the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant in FCA followed by two injections with 40 µg of the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

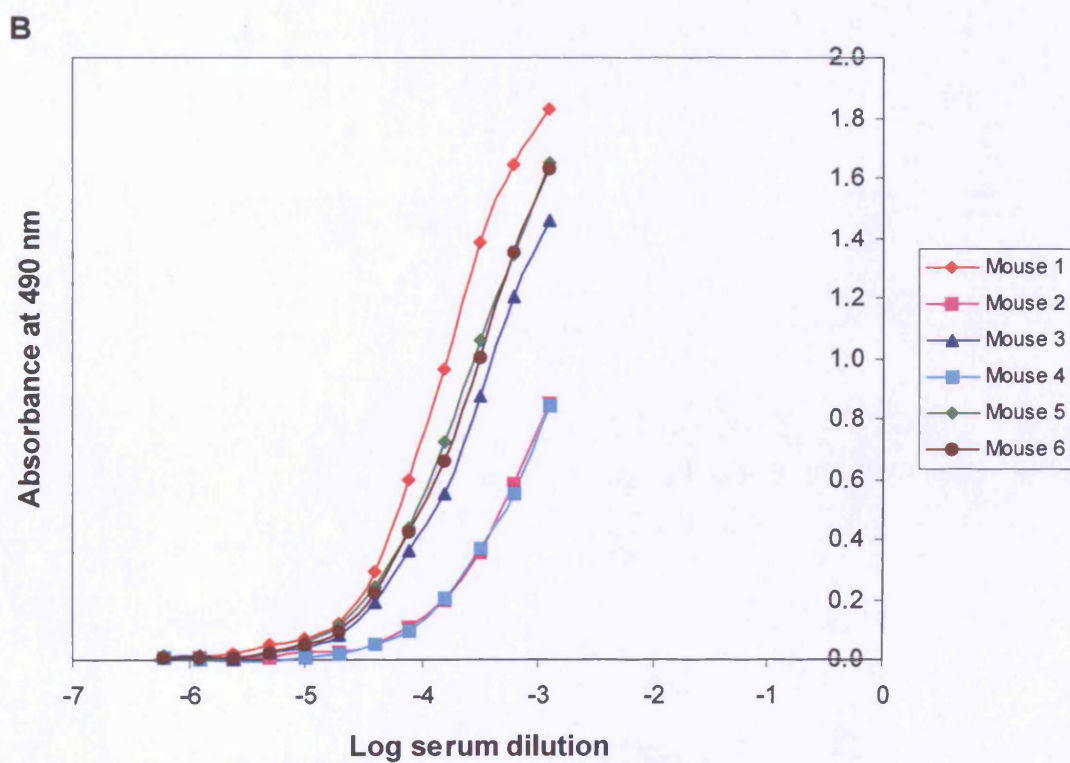
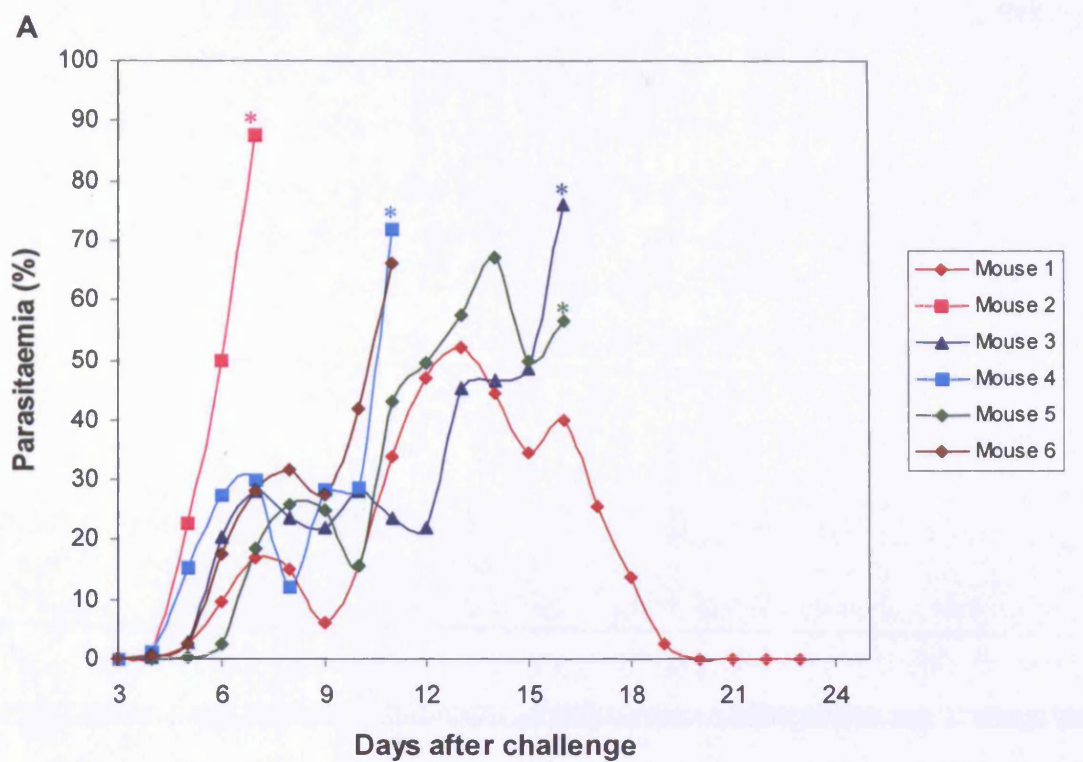


Figure 8.5: Course of *P. yoelii* infection in mice immunised with Glu28→Gln MSP1₁₉ variant and antibody binding curves.

A: Six BALB/c mice were immunised with 10 µg of Glu28→Gln MSP1₁₉ variant in FCA followed by two injections with 40 µg of Glu28→Gln MSP1₁₉ variant in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The percentage parasitaemia for the individual mice in the group is plotted on the graph. Asterisks indicate when a mouse died or was killed by a schedule one method.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Gln MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

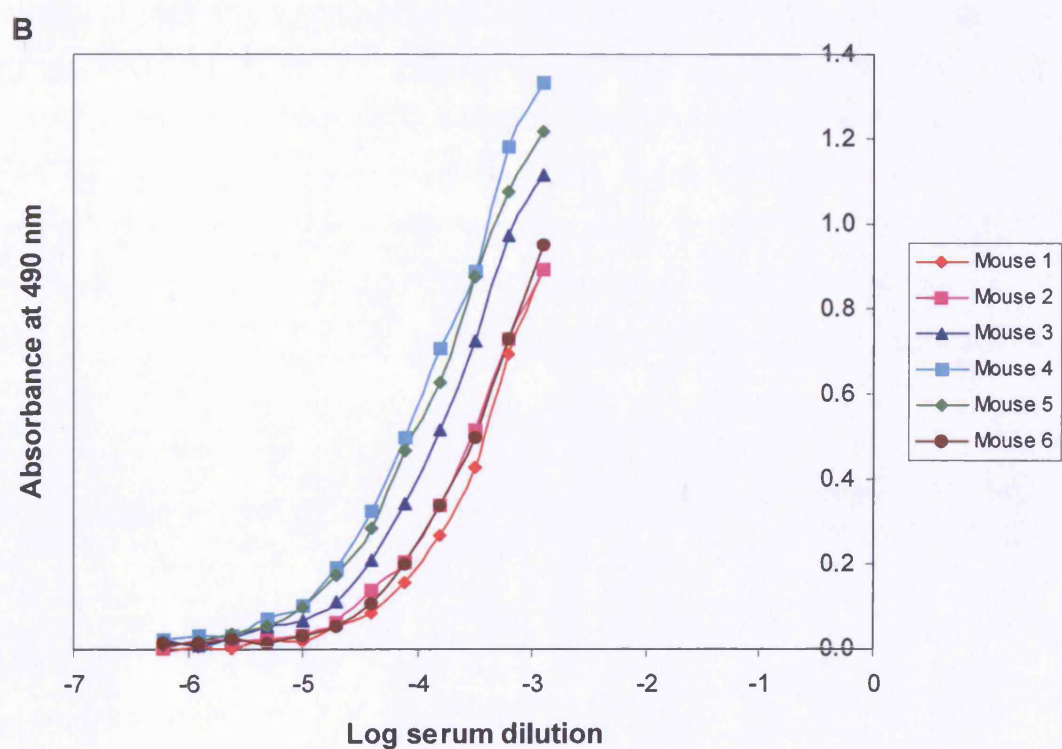
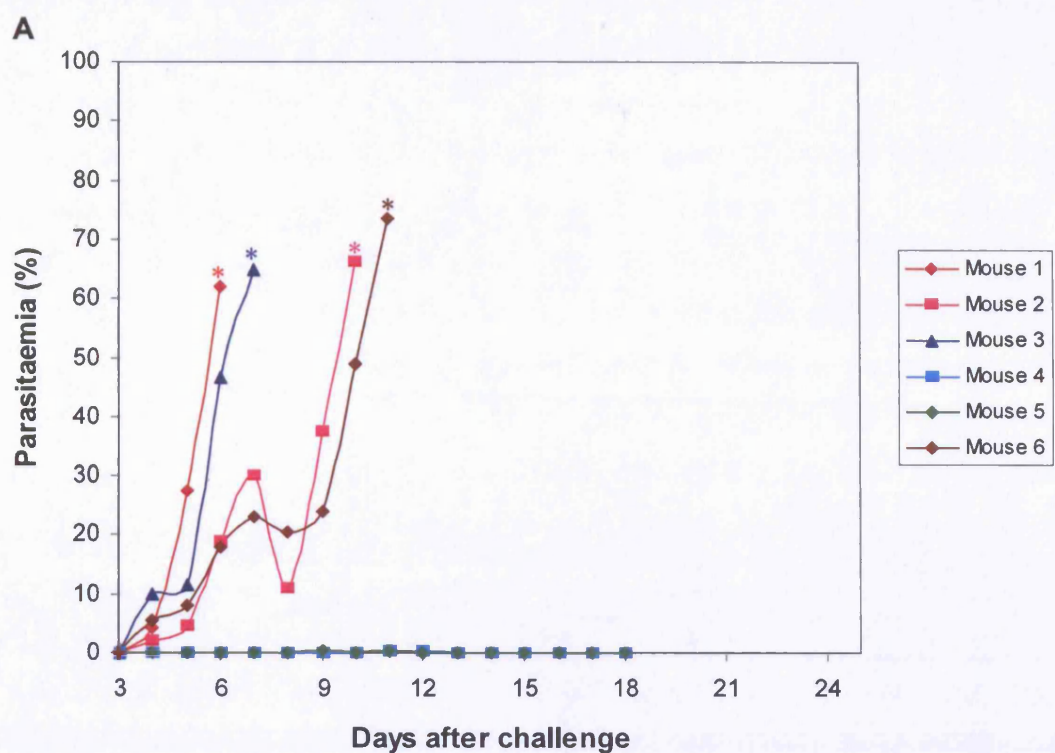


Figure 8.6: Course of *P. yoelii* YM infection in groups of mice immunised with wildtype and residue 28 MSP1₁₉ variants and antibody binding curves.

A: Groups of six BALB/c mice were immunised with 10 µg wildtype or MSP1₁₉ variants or GST in FCA followed by two injections with 40 µg of protein in FIA 21 and 42 days later. The mice were challenged 15 days after the final immunisation with 5×10^3 *P. yoelii* YM parasitized erythrocytes. The parasitaemia was followed daily from day 3 on Giemsa stained blood films. The average parasitaemia for the groups is plotted on the graph. The average parasitaemia for the mice immunised with wildtype MSP1₁₉ is shown in red, with GST is shown in pink, with Glu28→Lys MSP1₁₉ variant is shown in blue, with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in orange and with Glu28→Gln MSP1₁₉ variant is shown in green.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with wildtype and residue 28 MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for the wildtype, variant protein and GST less PBS control are shown on the graph. The serum samples for the six mice in the groups were pooled together. Serum from mice immunised with wildtype MSP1₁₉ is shown in red, with GST is shown in pink, with Glu28→Lys MSP1₁₉ variant is shown in blue, with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in orange and with Glu28→Gln MSP1₁₉ variant is shown in green.

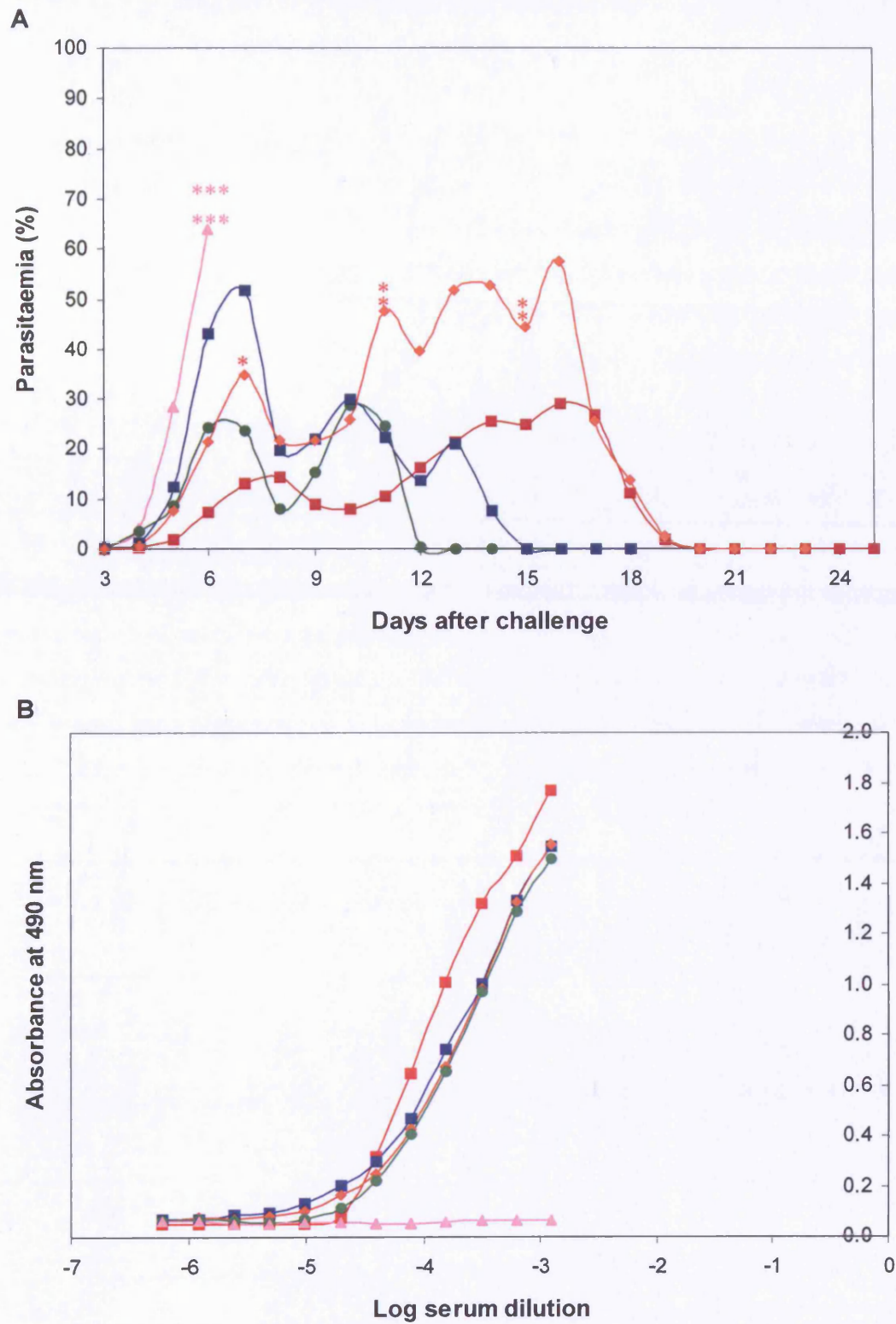


Figure 8.7: Antibody binding curves for serum from mice immunised with Glu28→Lys MSP1₁₉ variant against his-tagged wildtype and Glu28→Lys MSP1₁₉ variant.

A: 1 µg/ml of his-tagged Glu28→Lys MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ and his-tagged with Glu28→Lys MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph. Serum from mouse one is shown in red, from mouse two is shown in pink, from mouse three is shown in dark blue, from mouse four is shown in cyan, from mouse five is shown in green and from mouse six is shown in brown. The antibody binding curves for the sera binding to his-tagged wildtype MSP1₁₉ is shown with filled in shapes and the antibody binding curves for the sera binding to his-tagged Glu28→Lys MSP1₁₉ variant is shown with outlined shapes.

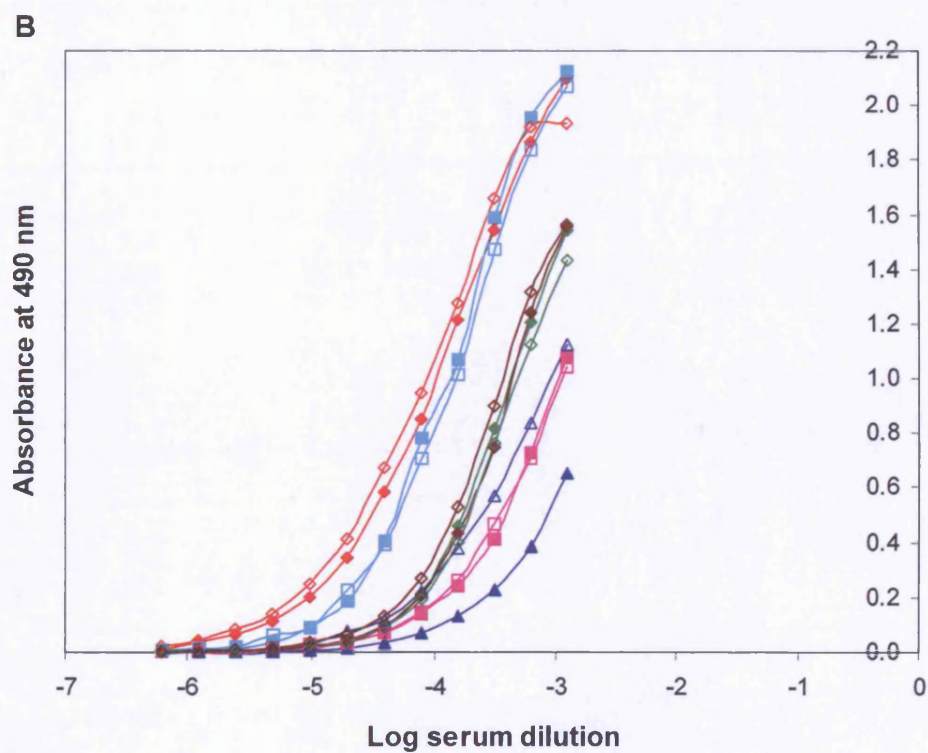
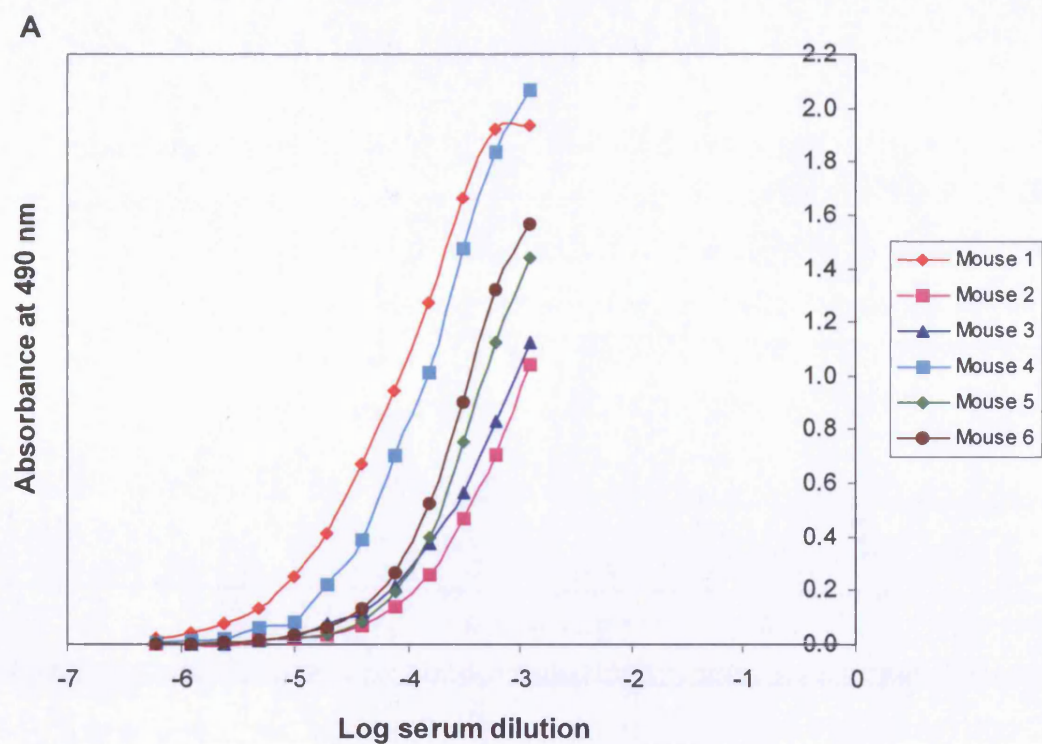


Figure 8.8: Antibody binding curves for serum from mice immunised with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant against his-tagged wildtype and double Lys16→Glu/Glu28→Lys MSP1₁₉ variant.

A: 1 µg/ml of his-tagged double Lys16→Glu/Glu28→Lys MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ and his-tagged double Lys16→Glu/Glu28→Lys MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with double Lys16→Glu/Glu28→Lys MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph. Serum from mouse one is shown in red, from mouse two is shown in pink, from mouse three is shown in dark blue, from mouse four is shown in light blue, from mouse five is shown in green and from mouse six is shown in brown. The antibody binding curves for the sera binding to his-tagged wildtype MSP1₁₉ is shown with filled in shapes and the antibody binding curves for the sera binding to his-tagged double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown with outlined shapes.

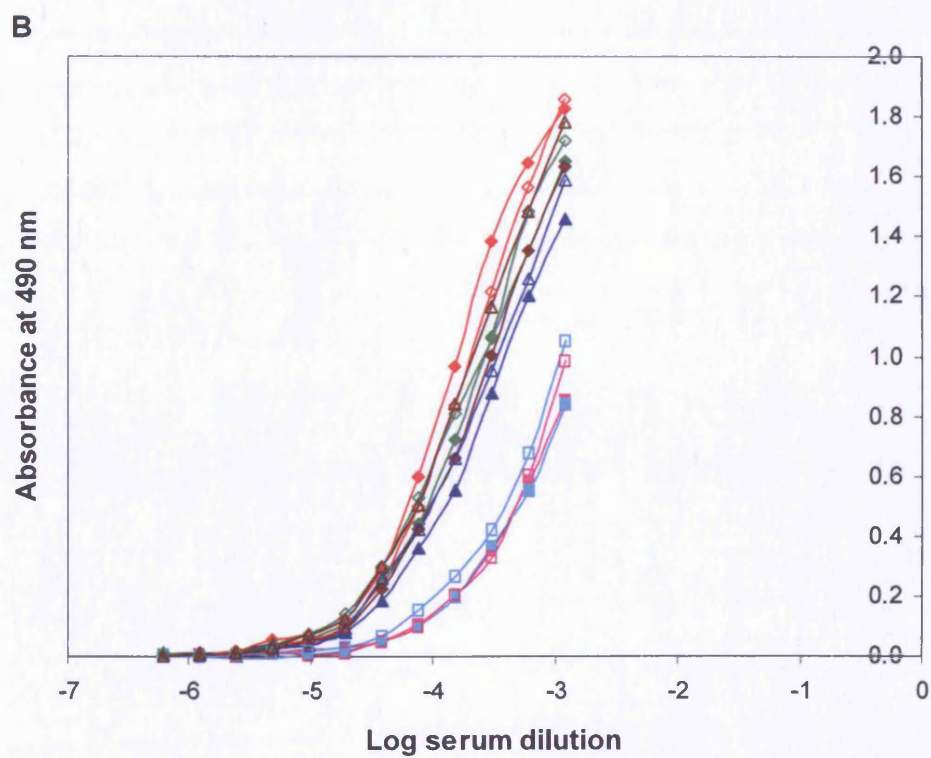
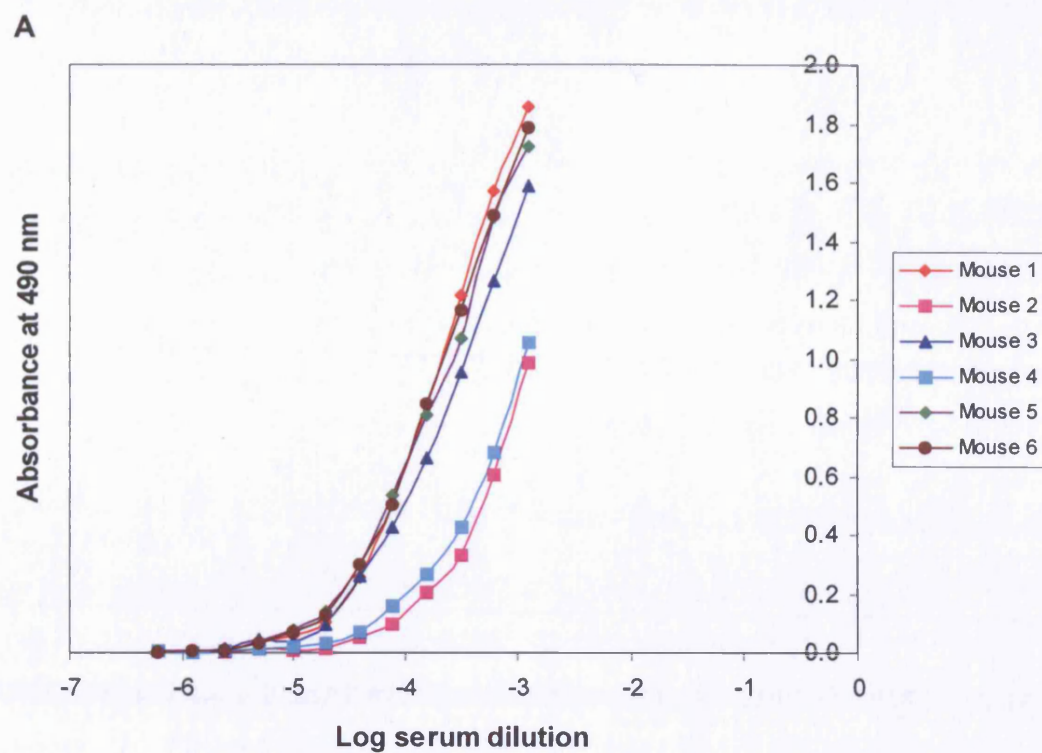
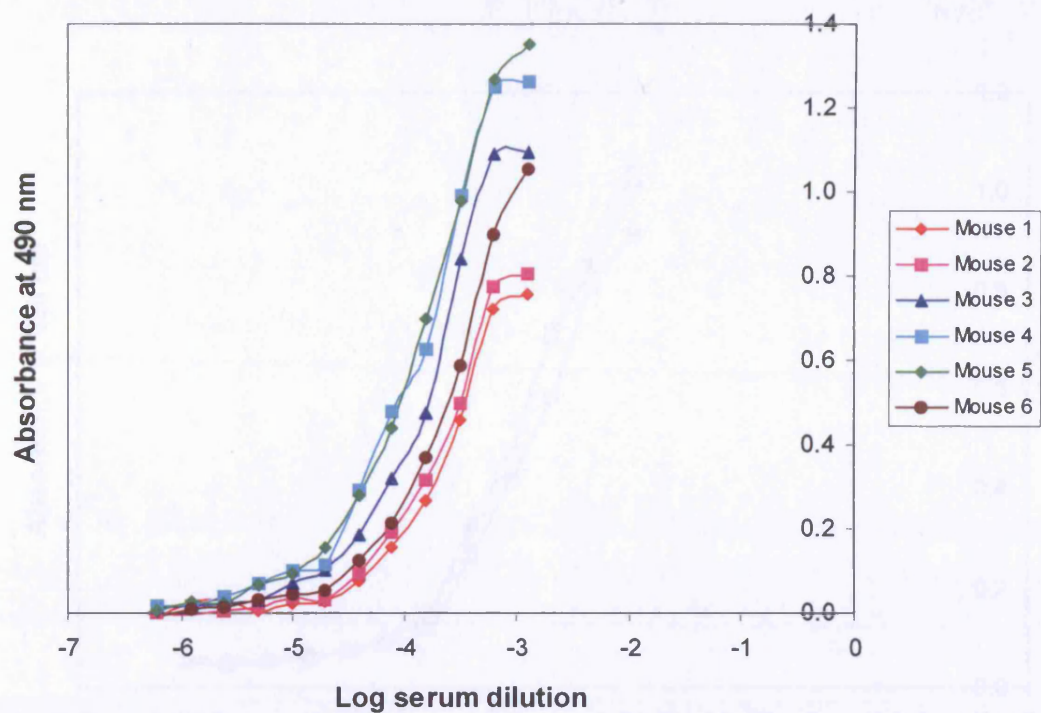


Figure 8.9: Antibody binding curves for serum from mice immunised with Glu28→Gln MSP1₁₉ variant against his-tagged wildtype and Glu28→Gln MSP1₁₉ variant.

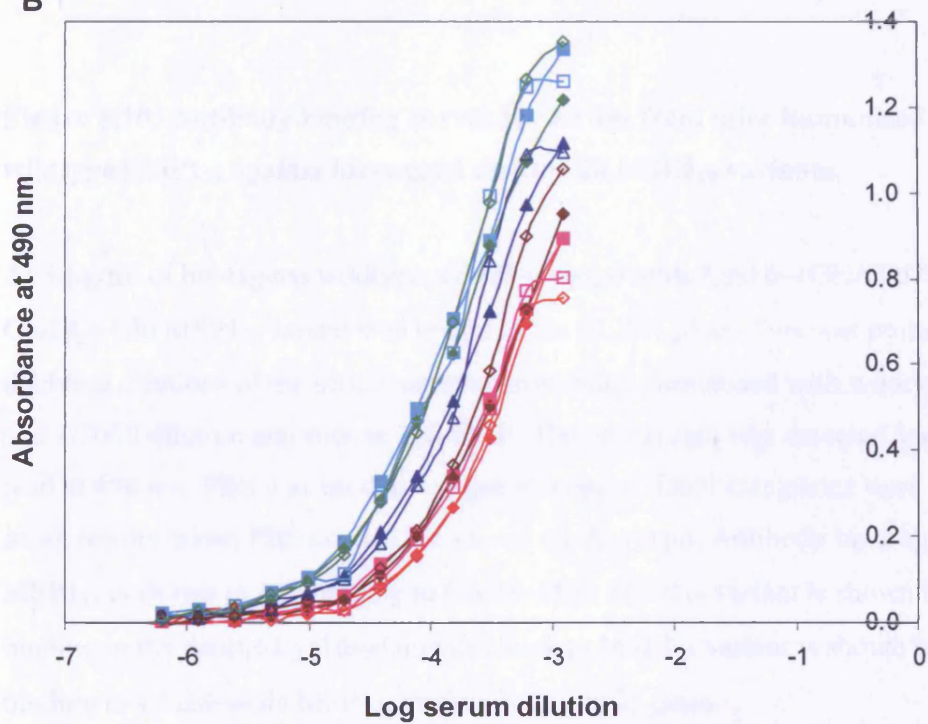
A: 1 µg/ml of his-tagged Glu28→Gln MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Gln MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph.

B: 1 µg/ml of his-tagged wildtype MSP1₁₉ and his-tagged with Glu28→Gln MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with Glu28→Gln MSP1₁₉ variant and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used. The mean results for each individual mouse less PBS control are shown on the graph. Serum from mouse one is shown in red, from mouse two is shown in pink, from mouse three is shown in dark blue, from mouse four is shown in cyan, from mouse five is shown in green and from mouse six is shown in brown. The antibody binding curves for the sera binding to his-tagged wildtype MSP1₁₉ is shown with filled in shapes and the antibody binding curves for the sera binding to his-tagged Glu28→Gln MSP1₁₉ variant is shown with outlined shapes.

A



B



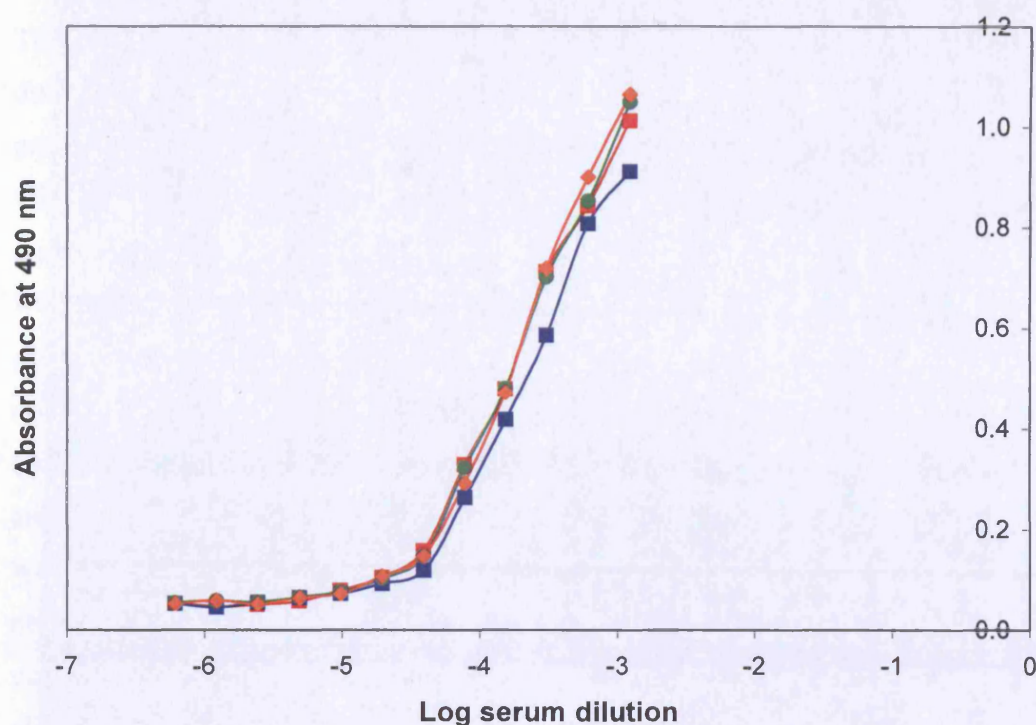


Figure 8.10: Antibody binding curves for serum from mice immunised with wildtype MSP1₁₉ against his-tagged residue 28 MSP1₁₉ variants.

A: 1 µg/ml of his-tagged wildtype, Glu28→Lys, double Lys16→Glu/Glu28→Lys and Glu28→Gln MSP1₁₉ variant was bound to the ELISA plate. This was probed with doubling dilutions of the serum samples from mice immunised with wildtype MSP1₁₉ and 1/2000 dilution anti-mouse IgG-HRP. The peroxidase was detected and absorbance read at 490 nm. PBS was used as a negative control. Duplicate plates were used and the mean results minus PBS control are shown on the graph. Antibody binding to wildtype MSP1₁₉ is shown in red, binding to Glu28→Lys MSP1₁₉ variant is shown in blue, binding to the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant is shown in orange and binding to Glu28→Gln MSP1₁₉ variant is shown in green.

8.5 Discussion

The immunisation studies presented in this chapter, have suggested that there are differences between the protection observed following immunisation with the wildtype protein and with the residue 28 variants. The overall results for the mice immunised with GST have shown that the GST-portion of the GST-MSP1₁₉ is not protecting the mice from challenge infection. The overall results have shown that immunisation with wildtype MSP1₁₉ protects against parasite challenge because five of the mice clear the parasite. Immunisation with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant does not protect against parasite challenge as all of the mice have high parasitaemia and only one mouse clears the parasites. The one mouse (mouse one) that had lower parasitaemia and cleared the parasites had an enlarged spleen and this could suggest that the parasites were being sequestered and that the parasitaemia may have been higher than what was observed from counting blood films. Immunisation with the Glu28→Lys MSP1₁₉ variant does not protect against parasite challenge as five of the mice were unable to clear the infection. The one mouse (mouse one) that had lower parasitaemia and was killed by a schedule one method on day 14 had malaria symptoms that were consistent with a higher parasitaemia than was counted from the blood films suggesting sequestration of parasites. The results from the immunisations with Glu28→Gln MSP1₁₉ variant were mixed. Four of the mice had a rapid increase in parasitaemia and were unable to clear the parasite. The results of the four mice suggested that the immunisations with Glu28→Gln MSP1₁₉ variant does not protect against parasite challenge. The other two mice however had very low parasitaemia with 0 % parasitaemia for a large proportion of the counts. These data would suggest protection by immunisation with Glu28→Gln MSP1₁₉ variant. There could have been error introduced into this group when the parasitized erythrocytes were administered to the mice with these two mice possibly getting less parasites or even no parasites.

The parasitaemia counts for the individual mice showed that there was variation between the individual mice in the group. This may have been because of variations in the experiment that could not be controlled. The mice used in the immunisation study were cousins because there were too many mice required for just brothers and sisters to

be used. This could have introduced some genetic variation between the mice which could have resulted in differences in immunity. Small differences in the amount of antigen the mice were immunised with could have occurred and this could have influenced the level of immune response. Differences in the uptake of the antigen between the individual mice following immunisation could have influenced the immune response produced. The behaviour of the individual mice may have influenced the progression of malaria, for example if a mouse was not eating or drinking as much as the other mice it could become unwell more quickly and be less able to fight the disease. The method of measuring the parasitaemia could also have introduced inaccuracies in the experiment because it relied on one person visually counting slides. If the mice were anaemic this could have led to inaccuracies in the counts because the blood smear from the anaemic mice did not give an even coverage of blood cells and cells appeared in clumps. This could have led to the blood cells that were counted to not be representative of the overall parasitaemia. The parasites could have sequestered in the spleen, brain or other organs and this would not be taken into account by counting the parasites in the blood and could result in lower counts and inaccuracies if variant proteins had altered the level of sequestration. For example, this may have occurred for mouse one in the group immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant as this mouse had an enlarged spleen. Additional errors could have been introduced in the decisions regarding when to kill the mice by a schedule one method as some mice became unwell at very low parasitaemia levels i.e. mouse one in the group immunised with Glu28→Lys MSP1₁₉ variant, while other mice were well at high parasitaemia i.e. mouse five in the group immunised with wildtype MSP1₁₉.

The ELISA results for the pooled serum samples show that there is no significant difference between the antibody titres to wildtype MSP1₁₉ produced by the wildtype and residue 28 MSP1₁₉ variant proteins. This suggests that the overall antibody titre may not be that important for protection against challenge infection. The ELISA results for the serum for the individual mice in the groups suggested there was more variation in the antibody titres. For the group of mice immunised with the wildtype MSP1₁₉, the antibody titres were all very similar. The antibody titre for mouse two that behaved differently to the other mice in the group and had a steady increase in parasitaemia did not have a significantly lower antibody titre than the rest of the mice in the group. This

suggests that there is not a clear correlation between the antibody titre and the ability of the mouse to protect against subsequent parasite challenge. This disagrees with the immunisation studies of Ling *et al.* (Ling *et al.*, 1994) where the antibody response to the parasite was highest in the mice that were protected against parasite challenge. The immunisation studies for the mice immunised with the residue 28 variants showed that there was more of a spread of antibody titres. The mice with the highest antibody titres in the group were those that survived the longest. There was more variation in the mice with lower antibody titres in the groups with mice with very similar antibody titres behaving differently on parasite challenge. For example, in the group immunised with the double Lys16→Glu/Glu28→Lys MSP1₁₉ variant, mouse two and four had very similar antibody titres but mouse two had a very rapid increase in parasitaemia and was killed by a schedule one method on day 7 and mouse four had a slower increase in parasitaemia and was killed by a schedule one method later on day 11. This shows that the antibody titres cannot be used as a prediction of the protection or course of the infection. The differences between the antibody titres for the residue 28 variants binding to wildtype MSP1₁₉ and residue 28 variant MSP1₁₉ protein suggests that a portion of the antibodies made to MSP1₁₉ are made to an area that is altered in the residue 28 variants. The relatively small difference between the binding to wildtype and residue 28 variants suggests that the portion of antibodies made to the area altered by residue 28 variants is only a small portion of the antibodies made. This suggests that the majority of the antibody response is made to the rest of the molecule that is not affected by the residue 28 variation. This observation is confirmed by the ELISA studies in this chapter for the binding of the pooled wildtype sera to the wildtype and residue 28 variant proteins. These data showed that there was no significant difference (with only a small difference with Glu28→Lys variant protein) between the binding to wildtype MSP1₁₉ and the residue 28 MSP1₁₉ variants. The data suggests that the majority of the antibodies produced on immunisation with MSP1₁₉ do not protect against challenge infection but only a small proportion of the antibodies that are made have the fine specificity to protect against challenge infection.

Chapter 9: Solving the 3D structures of wildtype *P. yoelii* MSP1₁₉ and Glu28→Lys MSP1₁₉ variant using NMR

9.1 Introduction

In this chapter, I will describe the calculation of the structure of wildtype MSP1₁₉ using NMR analysis. The predictions of structural changes as a result of the variations to MSP1₁₉ and the ¹⁵N-HSQC NMR spectra discussed in chapter 6 showed that *in silico* variation of the residues in the model was unable to accurately predict the differences in protein structure. It is therefore important to have an accurate protein structure to understand the affects on the structure caused by the amino acid variations.

There were limitations in the creation of the homology model (discussed in chapter 6) which means that it may not accurately reflect the structure of *P. yoelii* MSP1₁₉. The main limitation for the accuracy was the sequence identities of the template structures that were used. The sequence identities of the templates are listed in table 9.1. This shows that the sequence identities were around 50 % or less which means that the model's reliability would be decreased compared to a model made with higher levels of sequence identity (Schwede *et al.*, 2003).

Table 9.1: Sequence identities of the template MSP1₁₉ sequences used in the creation of the homology model		
Species	PDB structures	Sequence identity (%)
<i>P. falciparum</i>	1ob1F (Pizarro <i>et al.</i> , 2003)	49.95
	1ob1C (Pizarro <i>et al.</i> , 2003)	49.95
	1cejA (Morgan <i>et al.</i> , 1999)	49.95
<i>P. cynomolgi</i>	1b9wA (Chitarra <i>et al.</i> , 1999)	46.55
<i>P. knowlesi</i>	1nliC (Garman <i>et al.</i> , 2003)	50.84

In this chapter I will also describe the calculation of Glu28→Lys MSP1₁₉ variant protein using NMR analysis. The change to residue 28 affected *in vitro* antibody binding (as discussed in chapter 3) and resulted in loss of protection in *in vivo* studies (as discussed in chapters 5 and 8). The 2D ¹⁵N-HSQC NMR analysis (as discussed in chapter 6) indicated that there was a large amount of structural perturbation to the first EGF domain of the protein as a result of the Glu28→Lys variation. The 2D NMR analysis could not explain why this change to residue 28 had caused the level of structural perturbation and what the actual changes to the structure were. The best way to confirm how the Glu28→Lys variation had affected the structure of the protein was to calculate the structure of this variant and compare it to the wildtype structure.

The protein structures were not calculated for the other MSP1₁₉ variant proteins because with the exception of the double Lys16→Glu/Glu28→Lys variation and Glu28→Gln variation there were very few NH peaks that had moved in the 2D ¹⁵N-HSQC NMR spectra suggesting very little structural perturbation. The Glu28→Lys variant also had the largest biological affect on the protein therefore understanding it would be more biologically significant than solving the structures of the other variants.

9.2 Expression and Purification of ¹³C/¹⁵N labelled his-MSP1₁₉

For the 3D NMR analysis, proteins labelled with ¹⁵N and ¹³C were required. The *Pichia pastoris* expression system (described in chapter 6) was used to make doubly labelled ¹⁵N/¹³C proteins. In order to make ¹⁵N and ¹³C doubly labelled protein, the ammonium sulphate in the culture medium was substituted with ¹⁵N labelled ammonium sulphate and the methanol was substituted with ¹³C labelled methanol. The yeast uses the methanol as the sole carbon source for protein expression. This means that by substituting the methanol with ¹³C labelled methanol all of the protein will contain ¹³C instead of ¹²C. Large scale expression of the wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ variant was carried out to produce the 5 mg of pure protein required for NMR analysis. Figure 9.1 shows the wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ variant proteins before purification. This confirmed that the yeast were able to produce the proteins with the addition of ¹³C labelled methanol.

9.3 Assigning NMR Spectra for wildtype and Glu28→Lys MSP1₁₉ variant

¹³C and ¹⁵N 3D NMR Spectra were acquired for wildtype and Glu28→Lys MSP1₁₉ variant as described in materials and methods (section 2.5). The spectra were assigned manually using Sparky (Goddard & Kneller) and Xeasy (Bartels *et al.*, 1995) software. Figure 9.2 shows how the NMR spectra were used to determine the protein structures.

9.3.1 *HNCACB and CBCACONH NMR spectra*

The first stage involved assigning the backbone atoms. HNCACB and CBCACONH spectra were used for sequential assignment of the amino acid residues and to assign the NH peaks on the ¹⁵N-HSQC spectra. The spectra are named after the magnetisation transfer. The magnetisation transfer for the HNCACB and CBCACONH are shown in figure 9.3 (panel A). The HNCACB spectrum correlates the amide proton (NH) with the C α and C β . In the HNCACB spectra there are 4 peaks for each NH as follows: C α of the preceding amino acid; C β of the preceding amino acid; C α of the amino acid the NH belongs to and C β of the amino acid the NH belongs to. A schematic representation of this spectrum is shown in figure 9.3 (panel B). The assignment of residues 55 to 65 for wildtype MSP1₁₉ is shown in figure 9.4. This shows a small portion of the HNCACB spectra split into strips for the NH chemical shifts. The C α assignments are shown in red and the C β assignments are shown in blue. For residues 57 and 58 there are only C α peaks as these residues are glycines and there is a gap in the spectra for residue 61 because this is a proline and prolines do not have an NH peak on the spectrum. The CBCACONH spectrum (shown in the schematic in figure 9.3, panel B) correlates the amide proton (NH) with the C α and C β through the C=O. This means that in this spectrum only peaks for the C α and C β of the preceding amino acid and not the amino acid residue itself are present. The CBCACONH spectrum was therefore used to help in the assignment of the HNCACB spectra as it indicated which peaks were from the preceding amino acid. The assignments from these two spectra were used to assign

the ¹⁵N-HSQC spectra as shown in figure 9.5. The peaks with no labels are the peaks corresponding to the histidine tag and the factor Xa cleavage site.

9.3.2 HCCCONH and HCCH-TOCSY NMR spectra

The second stage in assigning the spectra was to assign the side chain residues. HCCCONH and HCCH-TOCSY spectra were used to assign the side chain protons and carbons. The HCCCONH correlates the amide proton (NH) of the amino acid with the ¹H in the side chain of the preceding amino acid. The HCCCONH spectra were assigned using the NH assignments from the ¹⁵N-HSQC experiments and the HNCACB experiments. The pattern of the ¹H in the side chain is characteristic of the amino acid type and can help identify the amino acid. Figure 9.6, 9.7 and 9.8 show three strips for three different amino acids from the wildtype MSP1₁₉ HCCCONH spectra. Figure 9.6 shows the NH chemical shift for valine 93 with the peaks for the ¹H side chain resonances of the preceding amino acid glycine 92. The two H α are not chemically equivalent and therefore appear as two separate peaks on the spectra. Figure 9.7 shows the NH chemical shift for alanine 18 with the peaks for the ¹H side chain resonances of the preceding amino acid asparagine 17. Asparagine 17 is an example of an AMX spin system. This means that it will have a characteristic pattern of one H α and two H β peaks. Other AMX amino acids include: cysteine, aspartic acid, serine, histidine, phenylalanine, tyrosine and tryptophan. Glutamic acid, glutamine and methionine show a very similar peak pattern to AMX amino acids apart from two additional peaks for H γ and are called AM(PT)X amino acids. Figure 9.8 shows the NH chemical shift for isoleucine 78 and the ¹H side chain resonances of the preceding amino acid isoleucine 77. Unlike the peak pattern for AMX and AM(PT)X amino acids, isoleucine has a unique peak pattern consisting of 6 peaks.

The HCCCONH spectrum only allows assignment of the ¹H side chain resonances and not the ¹³C side chain resonances. The HCCCONH spectrum provided a stepping stone for assigning all the side chain resonances using the HCCH-TOCSY spectrum. The ¹H side chain assignments from the HCCCONH spectrum and the C α and C β assignments from the HNCACB spectrum were used as a starting point for assigning the HCCH-TOCSY spectrum. HCCH-TOCSY stands for ¹H-¹³C-¹³C-¹H total correlation

spectroscopy and is used for the assignment of ¹H and ¹³C resonances (Cavanagh *et al.*, 2007). The magnetisation transfer for this experiment is shown in figure 9.9 (panel A). In the spectrum there are peaks for the ¹H side chain resonances at the chemical shift for each of the ¹³C side chain atoms. A schematic representation of the peak pattern that was seen for an aspartic acid residue is shown in figure 9.9, panel B. For the aspartic acid residue, it has a Cα and Cβ, therefore the peaks of the ¹H side chain resonances are seen at two ¹³C chemical shifts (one set of peaks for Cα and one set of peaks for Cβ). The two Hβ are coupled to each other which means there will be two lines of peaks for the Cβ chemical shifts (as seen in figure 9.9, panel B). A ¹³C-HSQC spectrum was used to help in the assignment of the HCCH-TOCSY spectrum. In the ¹³C HSQC spectrum there is a peak for each proton directly bonded to the carbon i.e. Cα-Hα, Cβ-Hβ2 and Cβ-Hβ3. There are no peaks to the adjacent carbons i.e. Cα- Hβ2.

9.3.3 Determining distance restraints for wildtype and Glu28→Lys MSP1₁₉

The HNCACB, CBCACONH, HCCCONH and HCCH-TOCSY experiments showed interactions between ¹H, ¹⁵N and ¹³C atoms through covalent bonds. In order to calculate the 3D structure of the protein information about interactions through space were required. This was achieved by acquiring 3D ¹⁵N-HSQC-NOESY spectra and ¹³C-HSQC NOESY spectra for wildtype and Glu28→Lys MSP1₁₉ variant.

NOESY stands for Nuclear Overhauser Effect spectroscopy. The Nuclear Overhauser Effect or NOE is a through space interaction between protons that does not require through bond coupling. A 2D ¹H-¹H NOESY spectrum correlates hydrogen atoms that are less than 5 Å apart in space. On the NOESY spectrum cross peaks are therefore seen for hydrogen atoms that are less than 5 Å apart. The intensity of the peak is proportional to r^{-6} where r is the internuclear distance. For MSP1₁₉, 3D ¹⁵N-HSQC-NOESY spectra were acquired because 2D ¹H-¹H NOESY experiments would have resulted in overlap of the ¹H-¹H NOE cross peaks. The 3D ¹⁵N-HSQC NOESY experiment overcomes this by combining the ¹⁵N-HSQC and NOESY experiments. This means that the overlapped ¹H cross peaks are resolved over the chemical shift frequencies of the directly attached ¹⁵N. This is shown in the schematic representation in figure 9.10. The ¹³C-HSQC-

NOESY combines the ¹³C-HSQC and NOESY experiments. The ¹⁵N-HSQC-NOESY and ¹³C-HSQC-NOESY experiments give information about the conformation of the protein and protein folding. They can also confirm residue assignments because residues have NOEs to their own side chain hydrogens. α -helices and β -sheets can be identified in the NOESY spectra. The strips for residues 20 to 22 in the ¹⁵N-HSQC-NOESY spectrum of wildtype MSP1₁₉ are shown in figure 9.11. These residues form part of the β -sheet. In the figure, the NOEs between adjacent NHs are very small. This is because in a β -sheet they are far apart. In a β -sheet there are strong NOEs between the H α of residue *i* and the NH of residue *i* + 1. The α -helix shows a different pattern of NOEs to the β -sheet, with NOEs observed to residues that are 3 apart in the primary sequence. For an α -helix there are strong NOEs between NH of residue *i* and NH of residue *i* + 1 and between H β of residue *i* and NH of residue *i* + 1. Figure 9.12 shows the strips for residues 9 to 11 from the wildtype MSP1₁₉ ¹⁵N-HSQC-NOESY spectra, which are involved in a turn. The NHs are close together for all three residues therefore there are NOEs between all of them. For the ¹⁵N-HSQC-NOESY spectra and ¹³C-HSQC-NOESY spectra for the wildtype and Glu28→Lys MSP1₁₉ variant the NH was assigned for each residue and the NOE peaks were picked. The individual NOE peaks were not assigned manually but were assigned using ARIA to calculate the protein structures (as described in section 9.4.1).

9.3.4 Predicting Phi and Psi angles using TALOS

The chemical shifts of the backbone atoms of the protein are sensitive to the local conformation of the backbone. It has been observed that the difference between the chemical shifts observed in the spectra (called the secondary chemical shift) and the chemical shift expected if that residue was in a random coil conformation is correlated with the secondary structure of the protein. In order to predict the backbone angles of the protein to help calculate the structure TALOS was used. TALOS stands for Torsion Angle Likelihood Obtained from Shift and Sequence Similarity. TALOS is a database system of 186 proteins for the prediction of phi and psi torsion angles. In order to predict the backbone angles, TALOS uses the C α , C β , CO, H α and N chemical shift assignments for the protein and compares them to homologous proteins that have

similar secondary chemical shifts. TALOS provides a list of possible phi and psi angles for each residue and rates how good these predictions are (Cornilescu *et al.*, 1999).

Since TALOS predicts the phi and psi angles and is not an experimental measure there is the possibility of errors, therefore only TALOS predictions that were rated as good were used in the structural calculations.

9.3.5 D₂O exchange analysis

Hydrogen bonding (H-bonding) is important in the stability of a protein structure.

Information can be obtained on H-bonding in proteins by resuspending the protein in ²H₂O (D₂O) before NMR analysis. When proteins are dissolved in D₂O the NH and OH protons exchange with the D₂O to give ND and OD. When ¹⁵N-HSQC NMR data are acquired for a protein sample in D₂O, only NHs where the proton has not exchanged will be visible as a peak on the spectrum. NHs that have not exchanged are said to be protected. There are two ways in which the NH can be protected. In a large protein the NH can be protected if it is in the core of the protein and the solvent cannot get to this part of the protein. The NH can also be protected if it is involved in a hydrogen bond. When the protein is in solution the hydrogen bonds will be continuously breaking and re-forming. The faster the rate of breaking, the less strong the hydrogen bond and the more quickly the D₂O will be able to exchange with the NH. This means that the rate of NH exchange correlates with the strength of the hydrogen bonds. If the hydrogen bonds are very strong, it will take a very long time for the NH to exchange. If the hydrogen bonds are weaker, it will take less time for the NH to exchange. If the NH is not involved in any hydrogen bond or not buried in the centre of a very large protein the NH will exchange instantly with the D₂O. Time course studies running ¹⁵N-HSQC spectra with protein in D₂O can highlight the residues in the protein that are protected from D₂O exchange and therefore could be involved in hydrogen bonding.

D₂O exchange time courses were acquired for wildtype and Glu28→Lys MSP1₁₉ variant to identify potential hydrogen bonds and the strength of the hydrogen bonds. Comparing the D₂O exchange time course data for the two proteins can identify areas of structural differences where hydrogen bonds may have been broken in the Glu28→Lys MSP1₁₉ variant. The wildtype and Glu28→Lys MSP1₁₉ were resuspended in D₂O and ¹⁵N-HSQC spectra were acquired every 5 minutes for 2 hours at 25 °C. The NH peaks

were assigned based upon the ¹⁵N-HSQC spectra in water. The rate of exchange of the protons with deuterium was determined and plotted on the histogram. The histogram in figure 9.13 shows the rate of exchange of NH in the wildtype MSP1₁₉ in red and Glu28→Lys MSP1₁₉ variant in blue. For NHs that were exchanging very quickly i.e. in less than 5 minutes, a rate of exchange could not be accurately determined so these residues were given an arbitrary value of 0. For NHs that were exchanging very slowly i.e. in greater than 3000 minutes, a rate of exchange could not be accurately determined so these residues were given an arbitrary value of 3000. The histogram shows that there are clear differences in the rate of exchange of NH in the two proteins in the first EGF domain (residues 1 – 48). In particular there is a cluster of NHs between residues 7 – 10 where there is instant exchange in the Glu28→Lys MSP1₁₉ variant and slower exchange for the wildtype MSP1₁₉. There are also differences in the rate of exchange between residue 21 – 29 where there is a slower rate of exchange in the wildtype MSP1₁₉. There are fewer differences in the rate of exchange between the two proteins in the second EGF domain (residues 49 – 99). The rates of exchange for the two proteins are mapped onto the best energy wildtype MSP1₁₉ NMR structure in figure 9.14.

9.4 Calculating the 3D structure of wildtype and Glu28→Lys MSP1₁₉ variant

9.4.1 *ARIA*

ARIA was used to calculate the structure of wildtype and Glu28→Lys MSP1₁₉ variant as described in materials and methods (section 2.5.4). ARIA stands for Ambiguous Restraint for Iterative Assignment. The role of ARIA is to assign ambiguous NOEs during the structure calculation. It does this by using ambiguous distance restraints and employing an iterative assignment strategy. An overview of the operations performed by ARIA is shown in figure 9.15. ARIA uses the complete assignment of proton chemical shifts and a list of partially assigned NOEs to calculate a group of structures. ARIA assigns each NOESY spectra separately and merges the data. ARIA creates a list of ambiguous distance restraints from the NOE peaks based on the chemical shift coordinates of the peak compared to the assignment list. It uses the list of restraints to calculate a set of structures. ARIA goes through eight iterations to improve the energy of the structure and finishes with a water refinement step. When ARIA has calculated the structures it creates a list of ambiguous and unambiguous NOEs and creates a list of peak violations that do not fit with the calculated structures. The violation list is checked manually to examine whether the violations were a result of an assignment error or if the peak is an artefact from water or a contaminant in the sample. In addition to the assignment list and NOESY spectra additional restraints can be put into ARIA for use in structural calculations. For wildtype and Glu28→Lys MSP1₁₉ variant the TALOS derived dihedral angle restraints were used in the calculations. A list of H-bonds was also added to the calculations as an iterative process by calculating the H-bonds (as described in section 2.5.4) and looking at the affect on subsequent ARIA calculations. Five disulphide bonds were also added to the structural calculations in an iterative process (as described in section 2.5.4). The final numbers of NOE distance restraints that were used in the final iteration of the ARIA structure calculations for wildtype and Glu28→Lys MSP1₁₉ variant are shown in table 9.2 and 9.3.

Table 9.2: ARIA NOE distance restraints used in the final structural calculation of wildtype MSP1₁₉

	intra-residue	sequential	medium range	long range
Unambiguous	1038.0	566.0	212.0	708.0
Ambiguous	103.8	151.9	109.9	341.3
Total	1141.8	717.9	321.9	1049.3

Table 9.3: ARIA NOE distance restraints used in the final structural calculation of Glu28→Lys MSP1₁₉

	intra-residue	sequential	medium range	long range
Unambiguous	845.0	457.0	179.0	667.0
Ambiguous	115.2	133.6	93.8	312.4
Total	960.2	590.6	278.8	979.4

Figure 9.16 shows the backbone traces of the twenty lowest energy structures for wildtype MSP1₁₉ shown in red and orange and for Glu28→Lys MSP1₁₉ variant shown in blue and cyan. The backbone traces for the group of wildtype structures show that the first EGF domain structures converge very closely and the second EGF domain structures do not converge as closely. The backbone traces for the group of Glu28→Lys MSP1₁₉ variant structures show that the first EGF domain structures also converge very closely and the second EGF domain structures do not converge as closely.

9.4.2 Evaluating the quality of the NMR structures

Procheck NMR was used to create Ramachandran plots to determine how good the wildtype and Glu28→Lys MSP1₁₉ variant structures were. The ramachandran plots for group of the twenty best energy structures for wildtype is shown in figure 9.17 and for Glu28→Lys MSP1₁₉ variant in figure 9.19. Figure 9.18 and 9.20 show the Ramachandran plots for the best energy structure for wildtype and Glu28→Lys respectively. All the Ramachandran plots show that all of the structures have over 90 %

of the residues with phi and psi angles in the most favoured and additional allowed regions for proteins.

9.5 Comparing the 3D NMR structure of wildtype *P. yoelii* MSP1₁₉ to the homology model

The best energy NMR structure of wildtype MSP1₁₉ was compared to the homology model (the creation of the homology model is discussed in chapter 6) using RasTop and MolMol. Figure 9.21 panels A and B show the backbone of the NMR structure of wildtype MSP1₁₉ and homology model respectively. The backbone structures were superimposed using residues 10 to 90 as shown in figure 9.21 panel C. Residue 1 to 10 and 90 to 99 were not used for superimposition because the very ends of the protein are not as well defined in the NMR structure as they would be moving around in the solution. Superimposing the backbone structures shows there are clear differences between the homology model and the NMR structure. The areas where there are significant changes in the backbone conformation are shown in the figure. The areas include residues 8 to 11 shown in blue and cyan; residue 12 shown in bright yellow and pale yellow; residues 70 to 75 shown in black and grey and residues 80 to 88 shown in dark green and light green. The comparison of the secondary structural elements of the NMR and homology model shown in figure 9.21 panel D suggests that the majority of the secondary structure is the same. The main difference between the homology model and NMR structure is an additional α -helix around residues 8 to 11. The other changes in the shape of the backbone seen in panel C correspond to loop areas.

9.6 Comparing the 3D NMR structure of wildtype and Glu28→Lys MSP1₁₉ variant

The best energy structure of the wildtype and Glu28→Lys MSP1₁₉ variant were compared using Insight II, MolMol, RasTop and Deep View/ Swiss PDB viewer to identify differences between the structures. The 2D ¹⁵N-HSQC NMR spectra shown in chapter 6 suggested that the majority of differences between the wildtype and Glu28→Lys MSP1₁₉ variant were in the first EGF domain. This suggests that the

comparison between the first EGF domain of the two structures will explain the effects of the residue 28 variation. The second EGF domain for Glu28→Lys MSP1₁₉ variant is not as well defined as the first EGF domain. This means that very small differences between the second EGF domain of wildtype and Glu28→Lys MSP1₁₉ variant may not be real. In the comparison between wildtype and Glu28→Lys MSP1₁₉ variant I will therefore focus on the first EGF domain and the interface between the two domains. Figure 9.22 shows the first EGF domain of the two proteins superimposed using residues 8 to 48. The residues at the ends of proteins are not as well defined because they are moving around in solution therefore residues 1 to 7 were not used for superimposing the structures. Superimposing the first EGF domain indicates there are differences between the two proteins in particular there is a difference in the shape of the loop consisting of residues 9 to 14. Figure 9.23 shows the residues where the NH peak in the 2D ¹⁵N-HSQC NMR spectrum had moved 0.2 ppm or more mapped onto the wildtype structure (panel A) and the Glu28→Lys MSP1₁₉ variant structure. The amino acid residues are coloured according to their properties: positively charged residues are shown in blue; negatively charged residues are shown in red, aromatic residues are shown in yellow and the other residues are shown in green. This shows that the side chain orientations of the residues have altered between the two proteins. Valine 9 has moved further out in the Glu28→Lys MSP1₁₉ variant. The other residues where there is a significant difference between the two proteins appear to be charged residues. The differences to the charged residues can be seen more clearly in figure 9.24 which shows only the charged residues where the NH moved more than 0.2 ppm in the 2D ¹⁵N-HSQC spectrum compared to the wildtype. This figure shows that arginine 12 and aspartic acid 13 appear to have swapped places between the wildtype and Glu28→Lys MSP1₁₉ variant with aspartic acid 13 moving towards lysine 28 and arginine 12 moving away. Arginine 31 has moved from pointing between the two EGF domains in the wildtype structure to being curved in Glu28→Lys MSP1₁₉ variant. Arginine 22 has also moved in the Glu28→Lys MSP1₁₉ variant, moving away from lysine 28. Aspartic acid 10 has moved away from arginine 12 towards arginine 22 and lysine 28.

The orientation of the two EGF domains relative to each other appears to be different between wildtype and Glu28→Lys MSP1₁₉ variant. In the interface between the domains there are four of the aromatic residues positioned very closely together. Figure

9.25 shows the aromatic residues in the two proteins. The residues are coloured according to the type of residue with phenylalanine shown in pink, tyrosine shown in green, tryptophan shown in purple and histidine shown in yellow. This shows that the orientation of the side chains of the aromatics at the interface between the two domains has changed between wildtype and Glu28→Lys MSP1₁₉ variant. The orientation of phenylalanine 21 and 94 has changed significantly between the wildtype and Glu28→Lys MSP1₁₉ variant. The orientation of tyrosine 89 has also changed.

Figure 9.1: NuPAGE gel analysis of expression of doubly labelled $^{13}\text{C}/^{15}\text{N}$ MSP1₁₉ proteins in *P. pastoris*.

2 × 500 ml cultures of GS115 cells containing multiple copies of the wildtype and Glu28→Lys MSP1₁₉ genes were grown in cultures containing ^{15}N labelled ammonium sulphate induced with ^{13}C labelled methanol to express the $^{13}\text{C}/^{15}\text{N}$ labelled proteins over 96 hours. 1 ml samples were removed every 24 hours. The supernatant was concentrated 10 times and run on pre-cast NuPAGE 12 % Bis-Tris polyacrylamide gels in MES buffer under reducing conditions and stained with Coomassie blue.

Lane 1 – molecular mass markers, lane 2 – Sample of supernatant from GS115 cells expressing wildtype MSP1₁₉ after 24 hours, lane 3 – GS115 cells expressing wildtype MSP1₁₉ after 48 hours, lane 4 – GS115 cells expressing wildtype MSP1₁₉ after 72 hours, lane 5 – GS115 cells expressing wildtype MSP1₁₉ after 96 hours, lane 6 – GS115 cells expressing Glu28→Lys MSP1₁₉ variant after 24 hours, lane 7 – GS115 cells expressing Glu28→Lys MSP1₁₉ variant after 48 hours, lane 8 – GS115 cells expressing Glu28→Lys MSP1₁₉ variant after 72 hours, lane 9 – GS115 cells expressing Glu28→Lys MSP1₁₉ variant after 96 hours .

The bands between 15 and 20 kDa are his-MSP1₁₉.

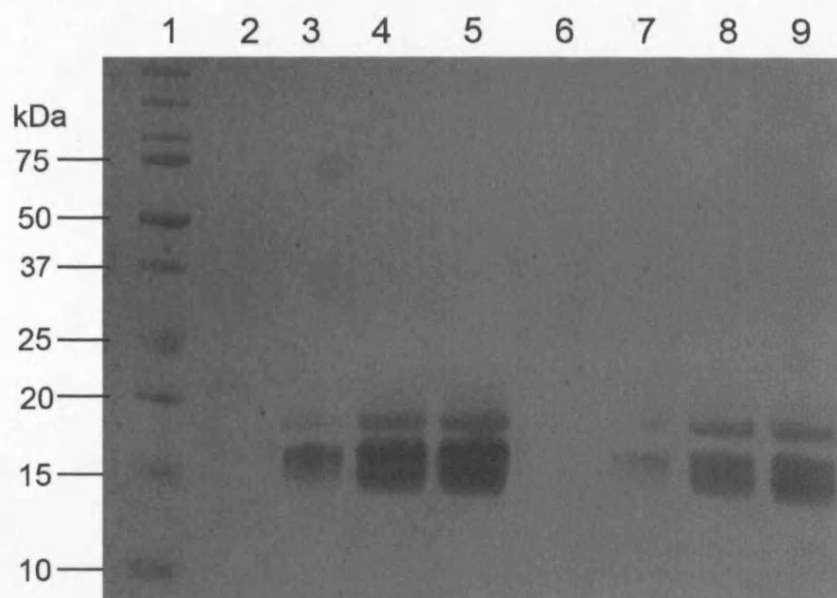


Figure 9.2: Schematic representation of how NMR spectroscopy was used for structural determination of wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ variants.

The steps are in black. The NMR experiments that were carried out for each step are in blue. The tools that were used to calculate restraints from the NMR data are in pink. The dashed arrows indicate where the data are used to check assignments and predictions in the spectra and then repeating the structural calculations.

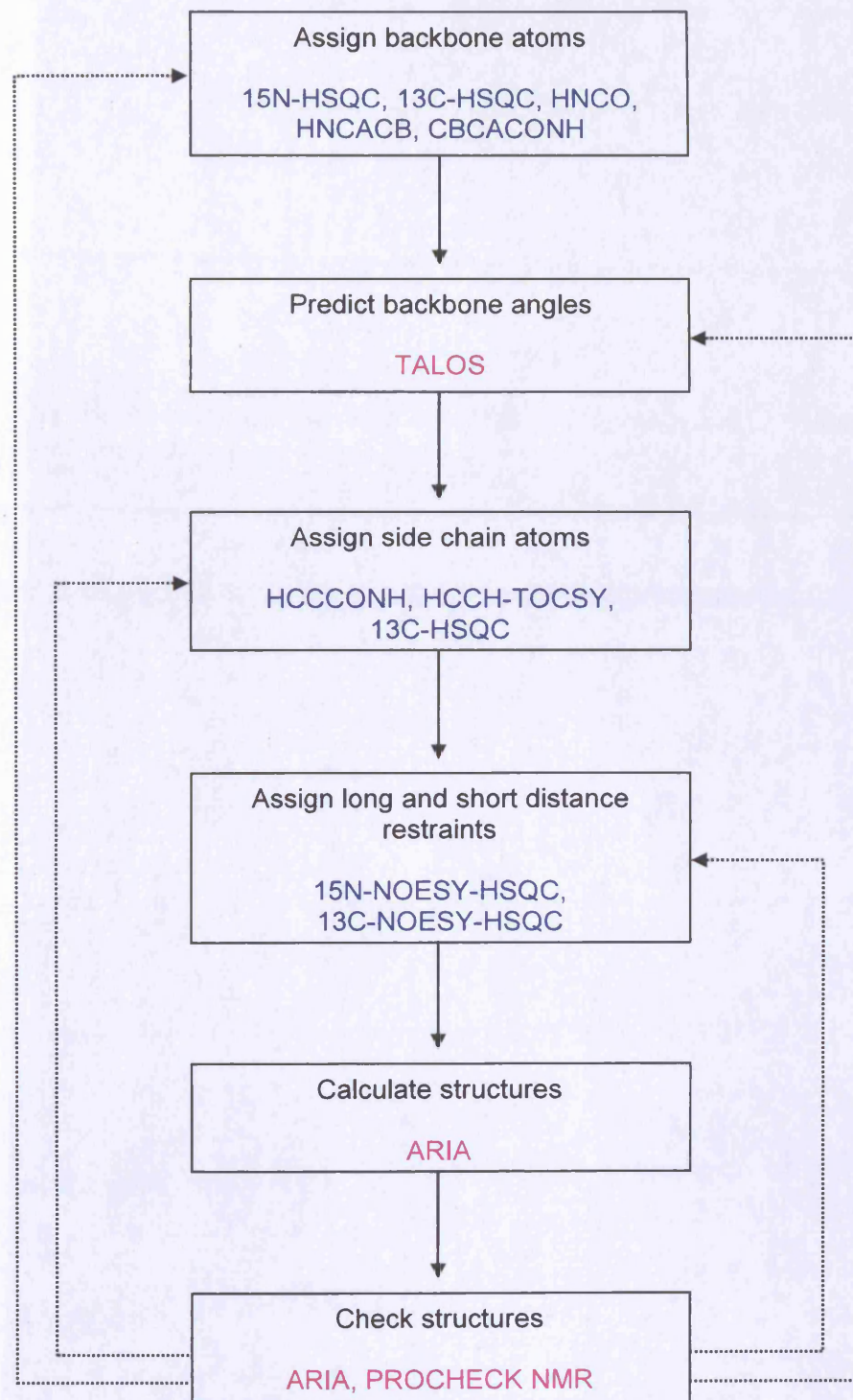


Figure 9.3: HNCACB and CBCACONH spectra.

A: The magnetisation transfer in the HNCACB and CBCACONH experiments. The magnetisation transfer is indicated by the blue arrows. In the HNCACB experiment, the magnetisation is transferred from the amide NH of residue i to the $C\alpha$ (CA) of the preceding residue ($i-1$) and then to the $C\beta$ (CB). The magnetisation is also transfer from the NH of residue i to the CA and then CB of residue i . In the CBCACONH experiment the magnetisation must go through the $C=O$, this means that the magnetisation is transferred from the CB of residue $i-1$ to the CA of $i-1$ and then through the $C=O$ to the NH of residue i . The magnetisation does not transfer to the CA and CB of residue i . Residue i is in red and $i-1$ is in black.

B: This shows a schematic representation of three strips for the HNCACB and CBCACONH of three sequential residues. The peaks for CA are shown in blue, the peaks for CB are shown in red. The peaks corresponding to CA of residue i and $i-1$ are labelled. The red dashed line indicates the connection between the CB of the neighbouring residues and the blue dashed line indicates the connection between the CA of the neighbouring residues.

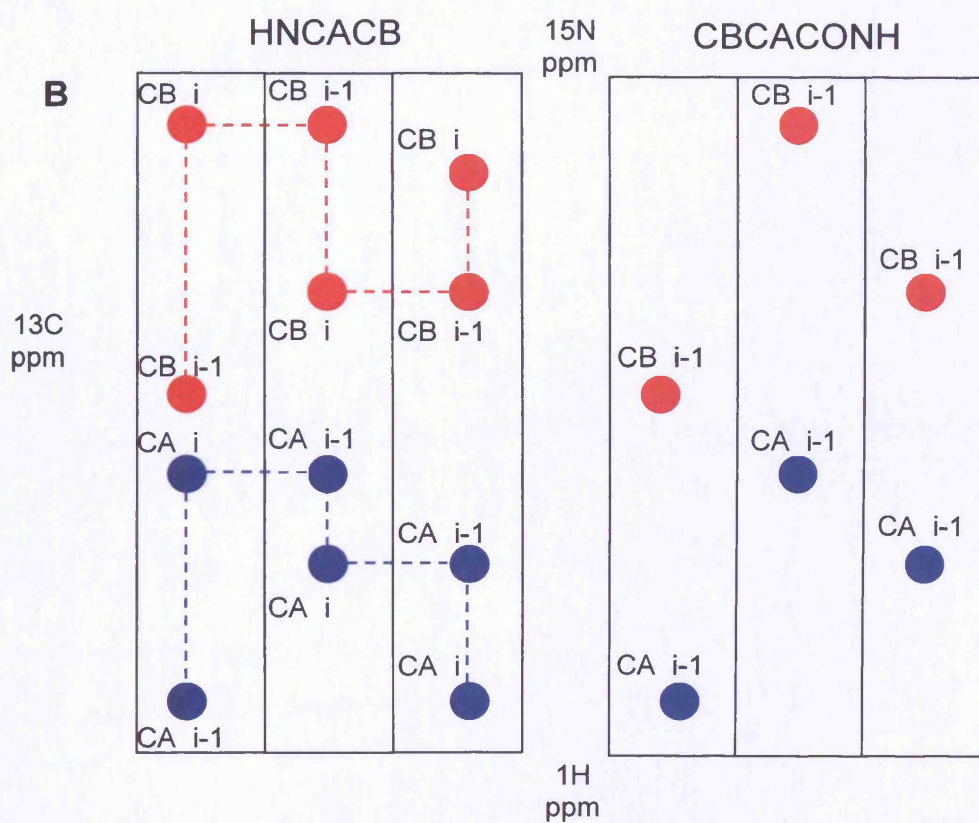
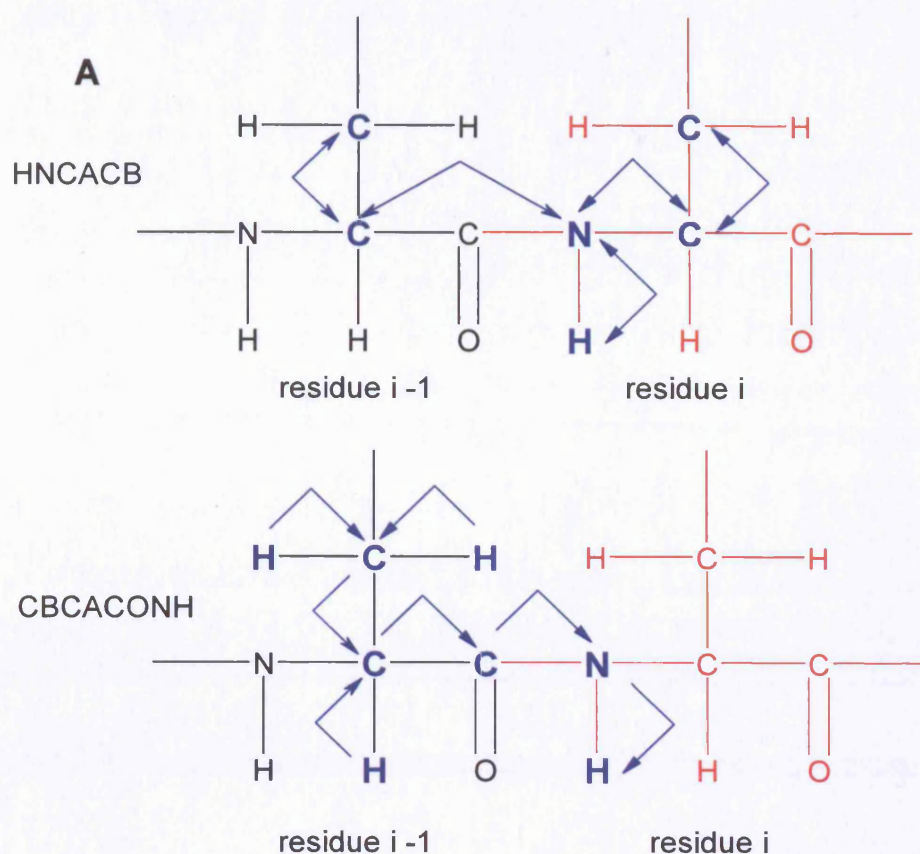


Figure 9.4: Assignment of residues 55 to 65 in the HNCACB spectrum of wildtype MSP1₁₉.

This figure shows the strips from the HNCACB spectrum of wildtype MSP1₁₉ for residues 55 to 60 and 62 to 65. There is no strip for proline 61 because it does not have an NH peak. The spectrum shows the C α and C β peaks at the chemical shift for the amide proton of the residue. In each strip there are peaks for the C α and C β for the residue and peaks for the C α and C β for the preceding residue. The red and blue lines show how the HNCACB spectrum can be assigned by walking along the C α and C β of adjacent residues. The connections between C α are shown in blue and the connections between C β are shown in red.

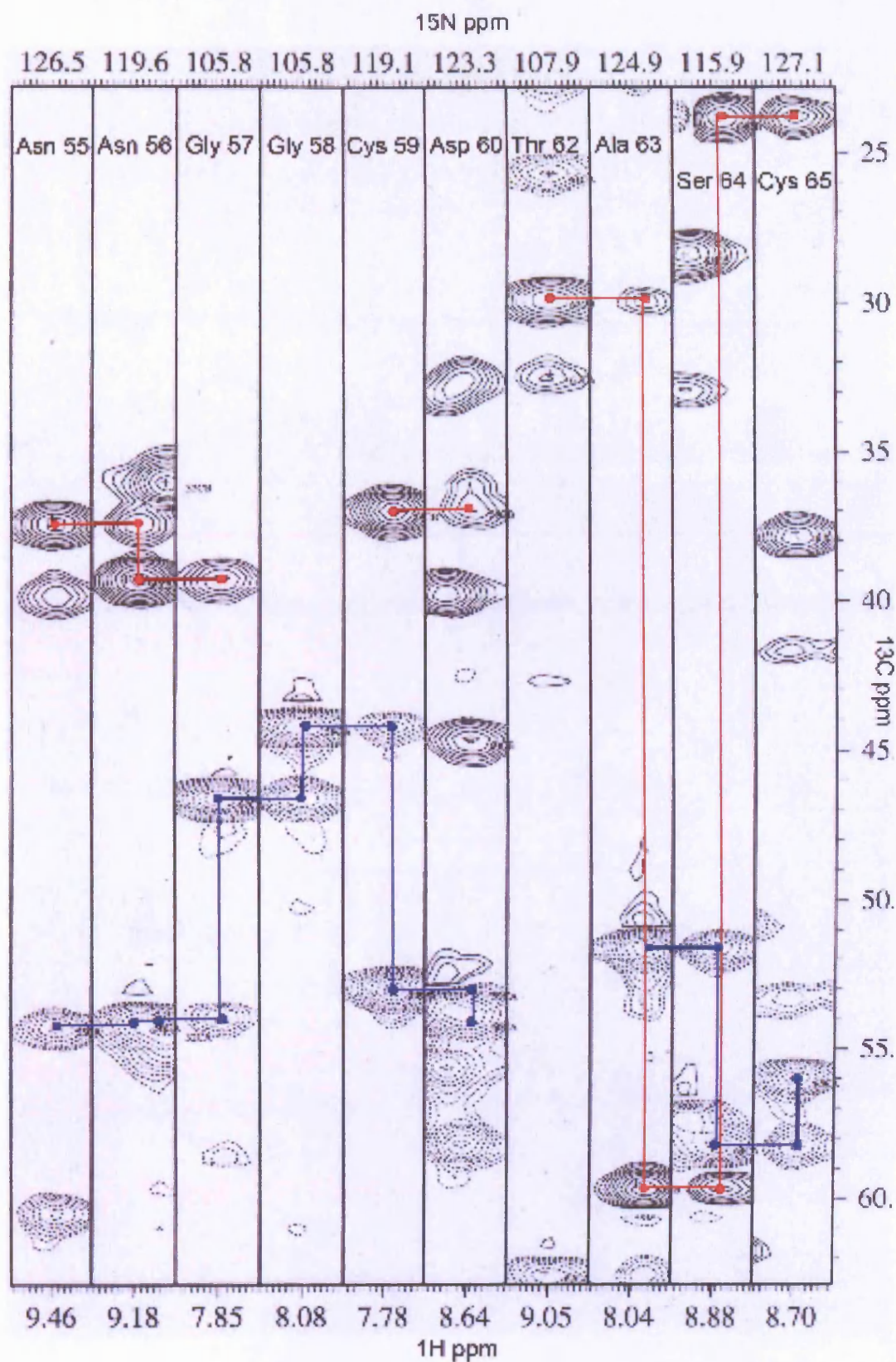
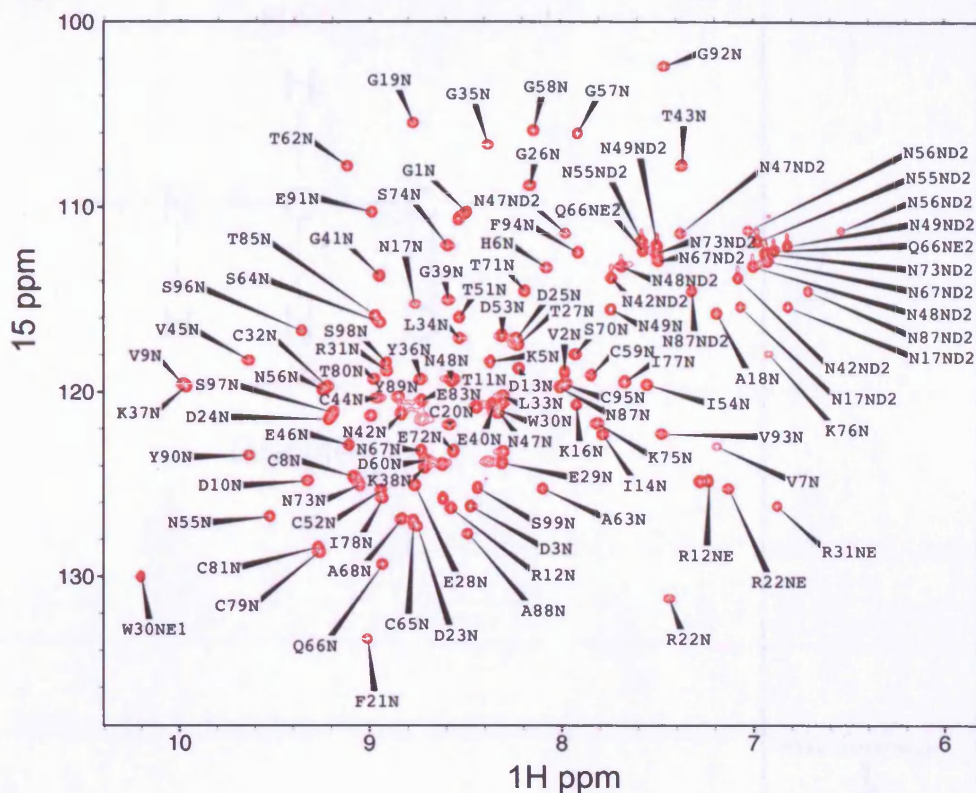


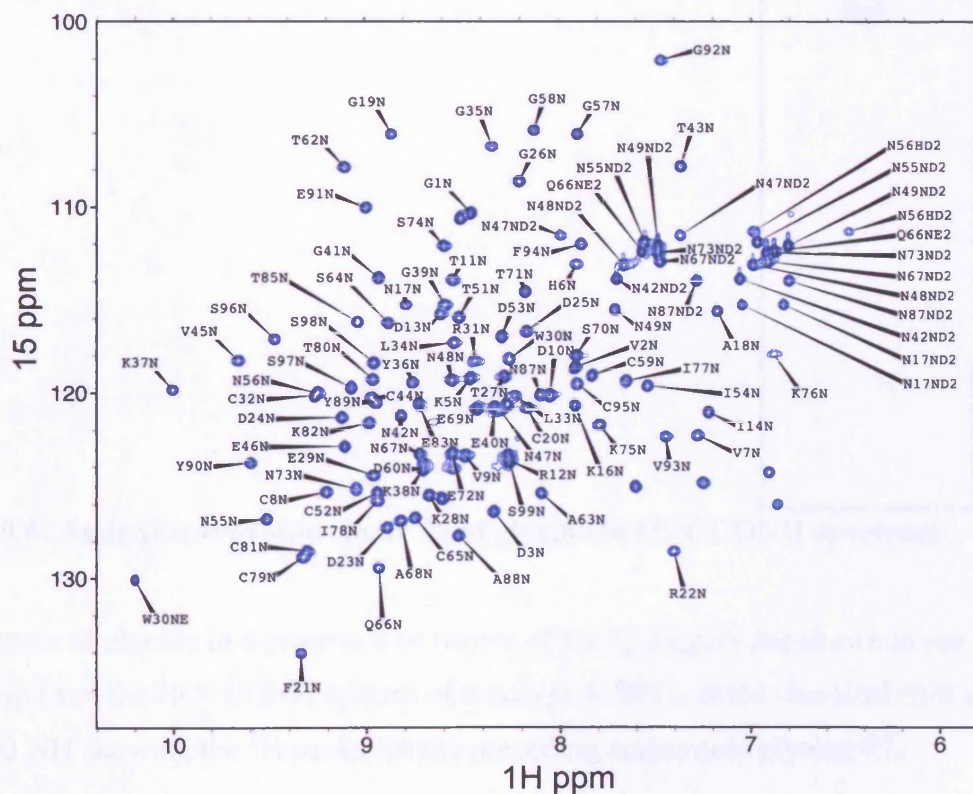
Figure 9.5: Assigned ^{15}N -HSQC spectra for wildtype and Glu28→Lys MSP1₁₉ variant.

Assignment of ^{15}N -HSQC Spectra of wildtype MSP1₁₉ (shown in red) and Glu28→Lys MSP1₁₉ (shown in blue) at 25 °C. The spectra were assigned using the data from the HNCACB, CBCACONH and ^{15}N -NOESY-HSQC NMR spectra. The unlabelled peaks are for the his-tag and the residues in the linker between the his-tag and the protein.

Wildtype



Glu28→lys



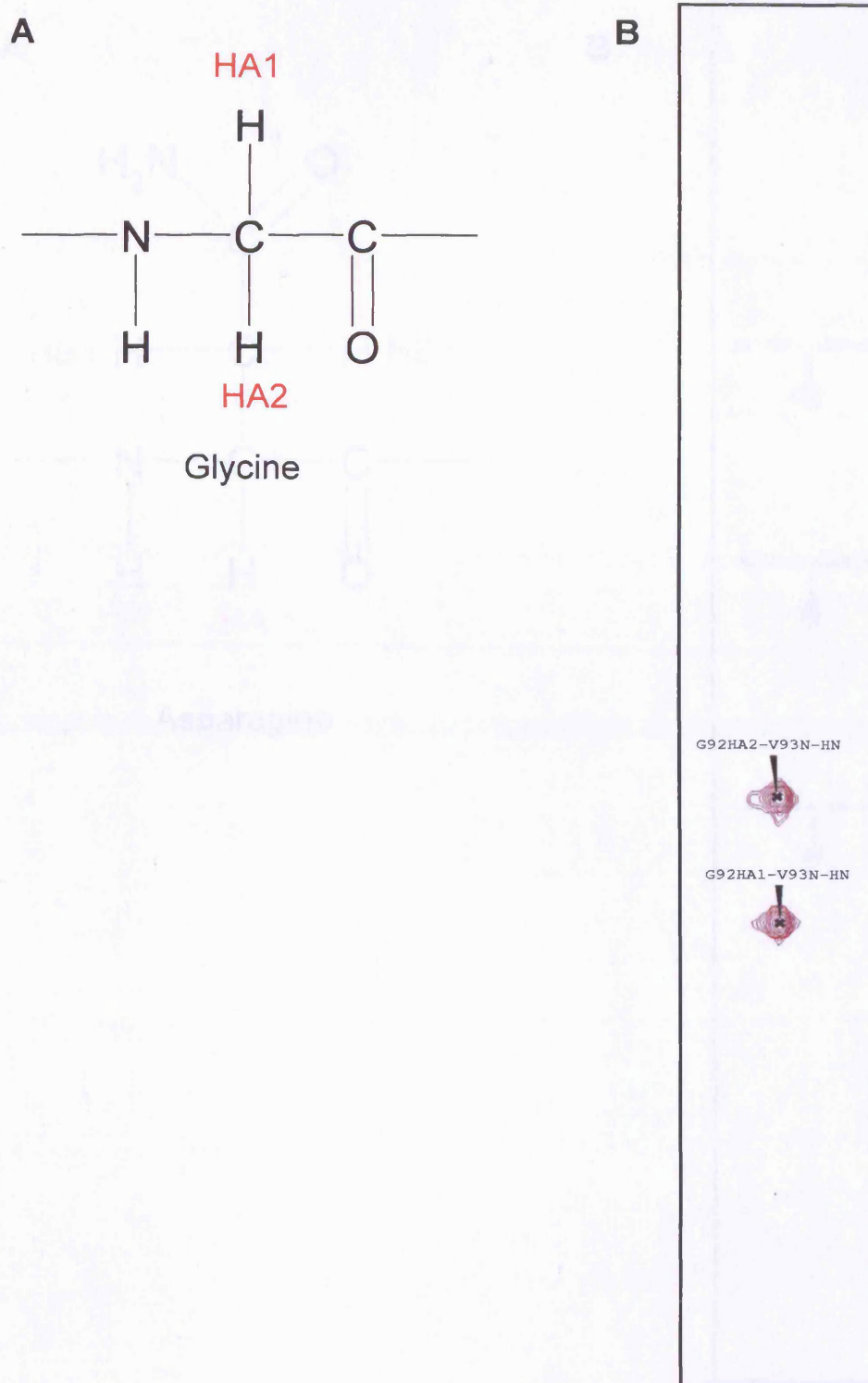


Figure 9.6: Assignment of side chain ^1H of glycine in HCCCONH spectrum.

A: Structure of glycine in a protein. The names of the hydrogens are shown in red.

B: A strip from the HCCCONH spectra of wildtype MSP1₁₉ at the chemical shift for valine 93 NH showing the ^1H peaks for the preceding amino acid glycine 92.

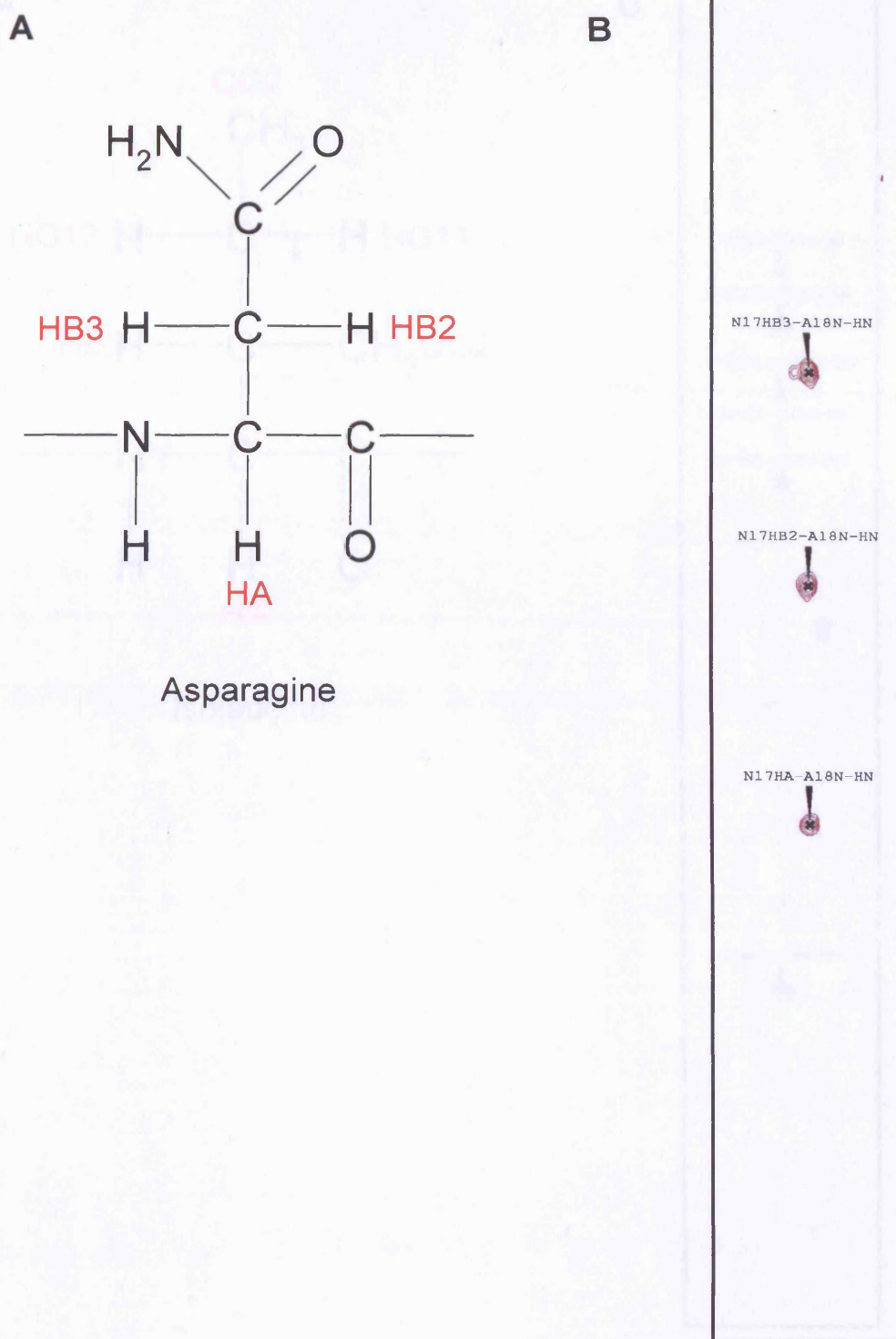


Figure 9.7: Assignment of side chain ^1H of asparagine in HCCCONH spectrum.

A: Structure of asparagine in a protein. The names of the hydrogens are shown in red.

B: A strip from the HCCCONH spectra of wildtype MSP1₁₉ at the chemical shift for alanine 18 NH showing the ^1H peaks for the preceding amino acid asparagine 17.

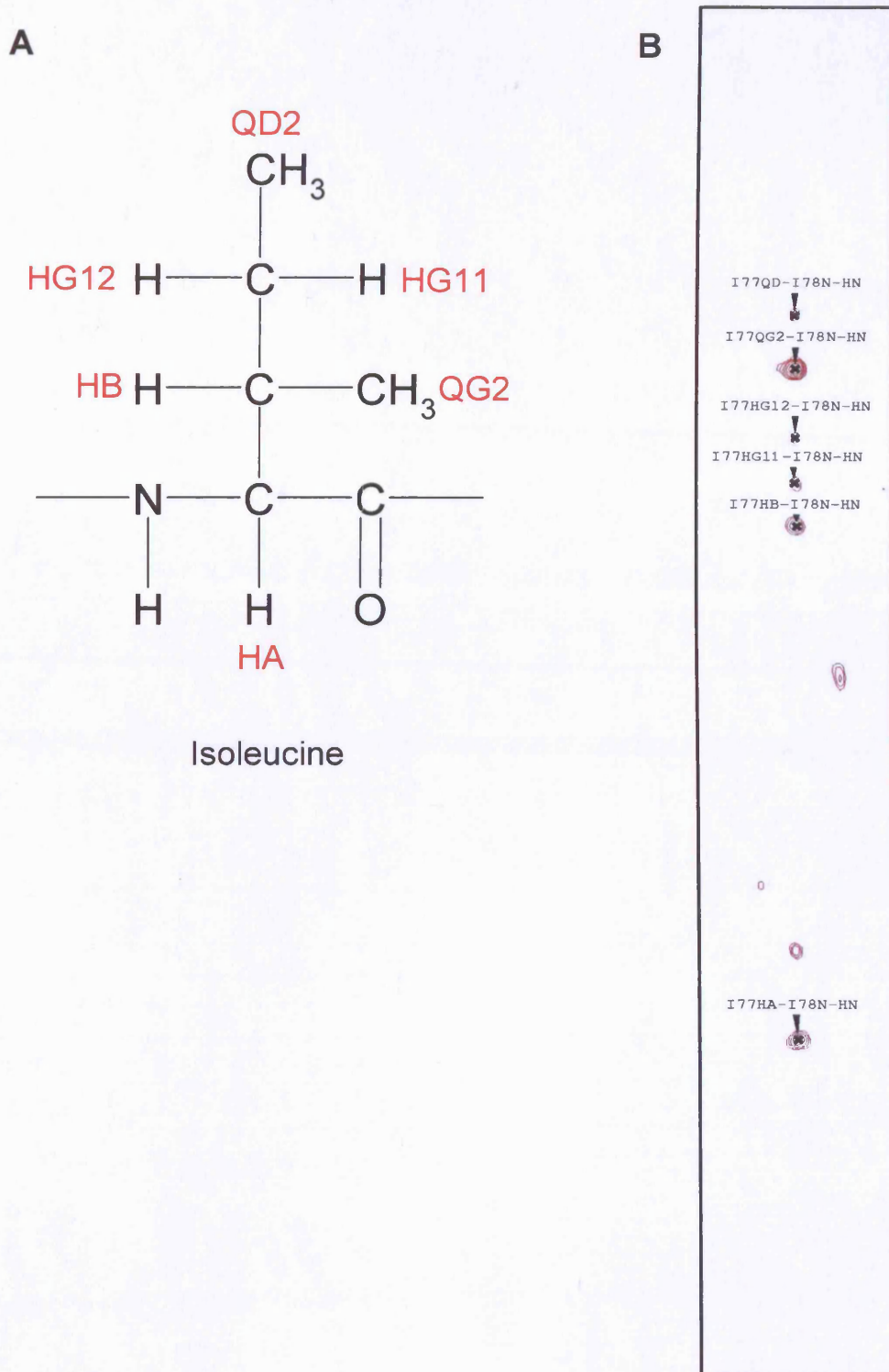


Figure 9.8: Assignment of side chain ¹H of isoleucine in HCCCONH spectrum.

A: Structure of isoleucine in a protein. The names of the hydrogens are shown in red.

B: A strip from the HCCCONH spectra of wildtype MSP1₁₉ at the chemical shift for isoleucine 78 showing the ¹H peaks for the preceding amino acid isoleucine 77.

Figure 9.9: HCCH-TOCSY NMR spectrum.

A: The magnetisation transfer in the HCCH-TOCSY experiment. The magnetisation transfer is indicated by the blue arrows. The magnetisation is transferred from the proton to the directed attached carbon atom, then to the neighbouring carbon atom and finally to the attached proton. Residue i is in red and $i - 1$ is in black.

B: The structure of aspartic acid in proteins. The protons are named in red and the carbons are named in pink.

C: A schematic representation of the HCCH-TOCSY spectra focusing on areas of the spectra in different planes that show the complete spin system of an aspartic acid residue. The top box represents the $C\alpha$ (CA) plane showing the peak for $H\alpha$ (HA) on the diagonal (represented by the dashed line) and the peaks for $H\beta 2$ and $H\beta 3$ (HB). The lower box represents the $C\beta$ (CB) plane showing two peaks on the diagonal (represented by the dashed line) for each of the HB protons and the cross peaks between them.

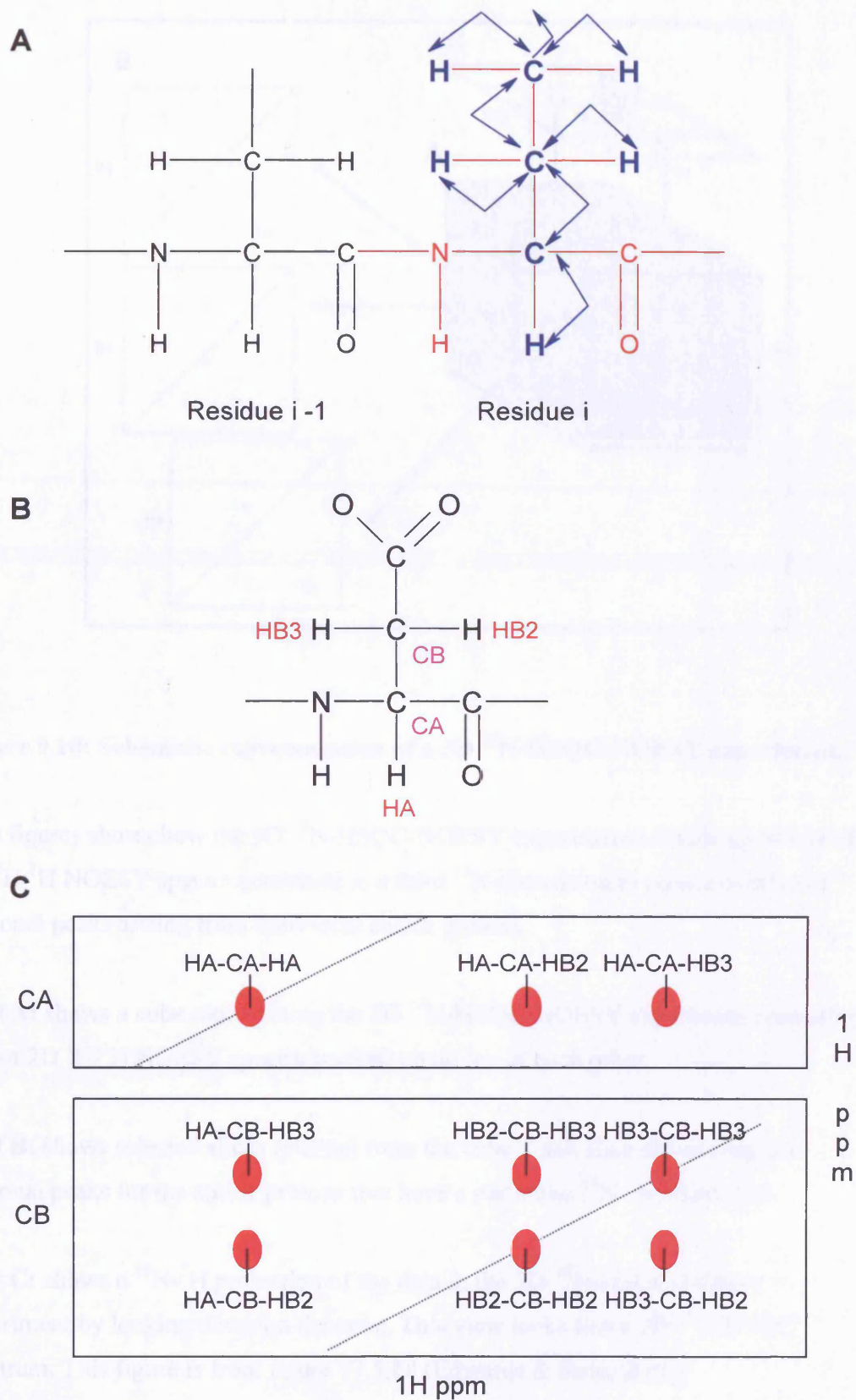


Figure 9.11: ^{15}N -HSQC NOESY spectrum for wildtype MSP1₁₉ residues 20 to 22 highlighting the NOEs between residues in a β -sheet.

Panel A: shows the arrangement of adjacent residues in a β -sheet. The blue arrow shows the NOE between the $\text{H}\alpha$ (HA) of residue i and the NH of residue $i + 1$.

Panel B: shows the strips for the NH ^{15}N chemical shift for residues 20, 21 and 22 which are part of a β -sheet. The blue lines indicate the NOEs between the HA of adjacent residues. The black line indicates the NOE between the NH of residue 20 and 21.

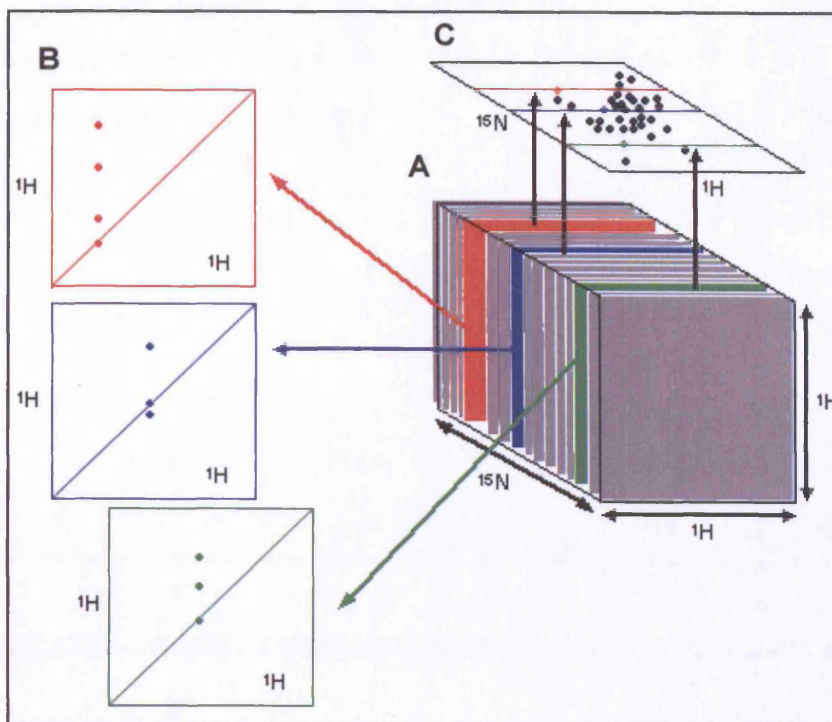


Figure 9.10: Schematic representation of a 3D ^{15}N -HSQC-NOESY experiment.

This figure shows how the 3D ^{15}N -HSQC-NOESY experiment is made up of lots of 2D $^1\text{H}/^1\text{H}$ NOESY spectra combined in a third ^{15}N -dimension to reduce overlap of diagonal peaks arising from individual amide protons.

Part A: shows a cube representing the 3D ^{15}N -HSQC-NOESY experiment consisting of lots of 2D $^1\text{H}/^1\text{H}$ NOESY spectra stacked up on top of each other.

Part B: shows selected slices (planes) from the cube. Each slice shows only the diagonal peaks for the amide protons that have a particular ^{15}N chemical shift.

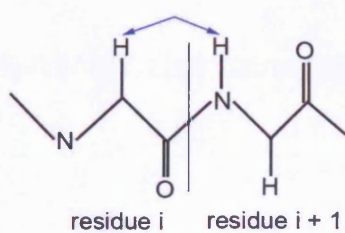
Part C: shows a ^{15}N - ^1H projection of the data in the 3D- ^{15}N -HSQC-NOESY experiment by looking down on the cube. This view looks like a 2D ^{15}N -HSQC spectrum. This figure is from figure 17.5.18 (Edwards & Reid, 2000).

Figure 9.12: ^{15}N -HSQC NOESY spectrum for wildtype MSP1₁₉ residues 9 to 11 highlighting the NOEs between residues involved in a turn.

Panel A: shows the alignment of adjacent residues involved in a turn. The nitrogen atom is highlighted in blue and the hydrogen atom is highlighted in green. The hydrogen atom is on top of the nitrogen pointing out of the page. This shows that the NH is close for all three residues involved in the turn.

Panel B: shows the strips for the NH ^{15}N chemical shift for residues 9, 10 and 11 which are part of a turn. The blue lines indicate the NOEs between the NH of adjacent residues. This shows that there are strong NOEs between the NH of adjacent residues.

A



B

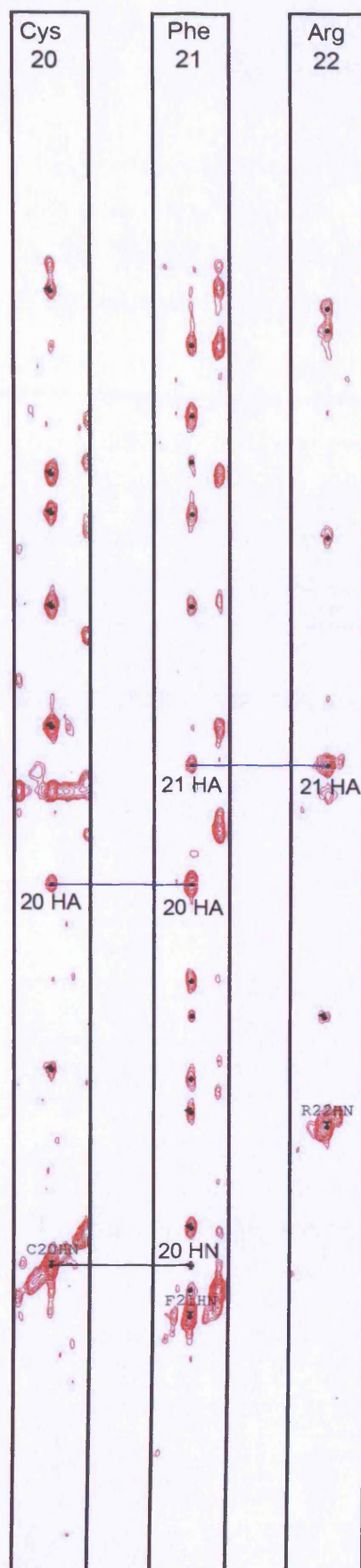


Figure 9.13: Histogram of D₂O exchange rates for wildtype and Glu28→Lys MSP1₁₉ variant.

The histogram shows the log D₂O exchange rates for wildtype MSP1₁₉ in red and Glu28→Lys MSP1₁₉ variant in blue. The D₂O exchange rates were calculated from a D₂O exchange time course where the proteins were dissolved in D₂O and ¹⁵N-HSQC were acquired every 5 minutes for 2 – 3 hours at 25 °C. The NHs that were exchanging with the D₂O in less than 5 minutes were given an arbitrary value of 0 and the NHs that were exchanging with the D₂O in more than 3000 minutes were given an arbitrary value of 3000. Prolines residues have been given a value of 0 as they have no NH to exchange.

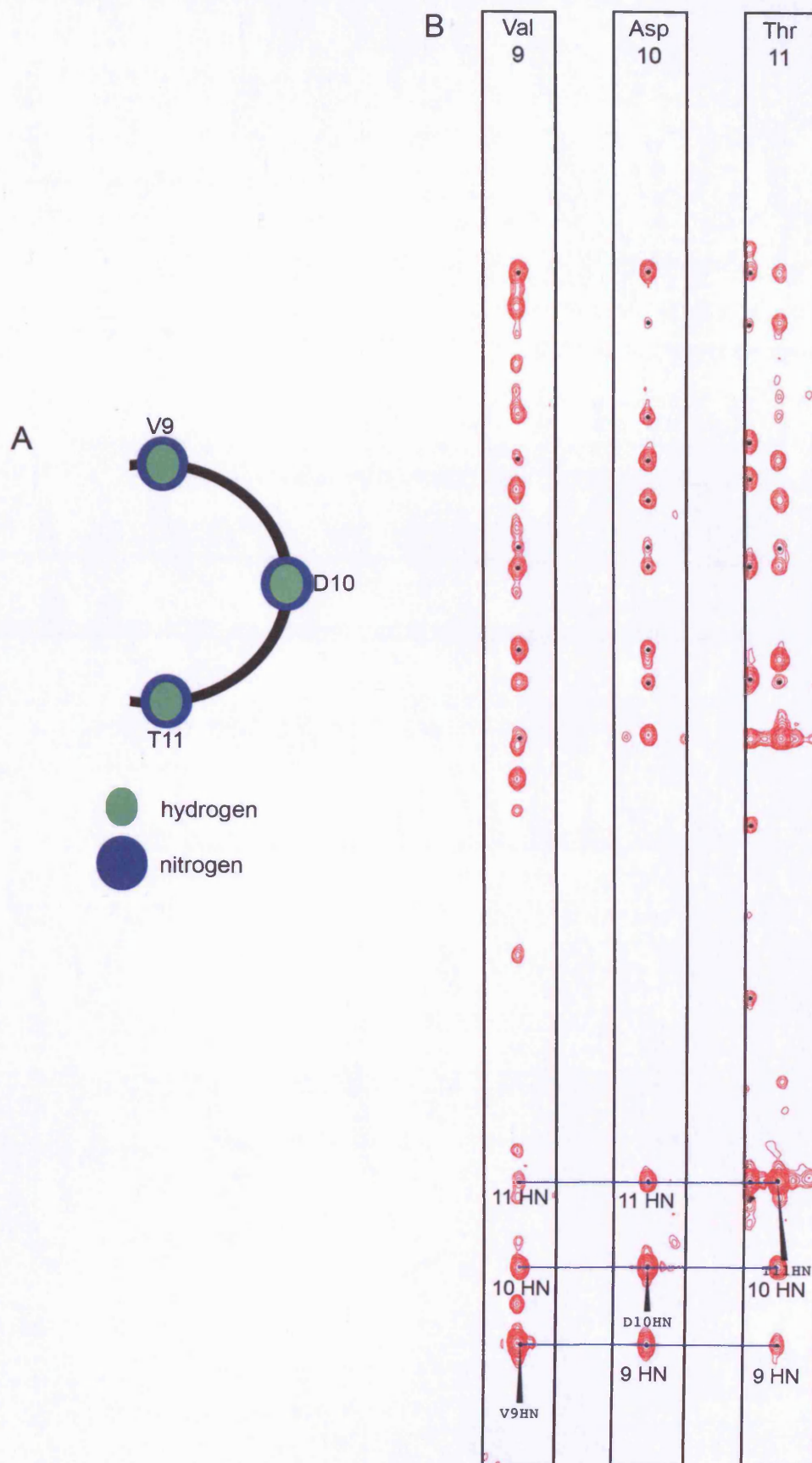
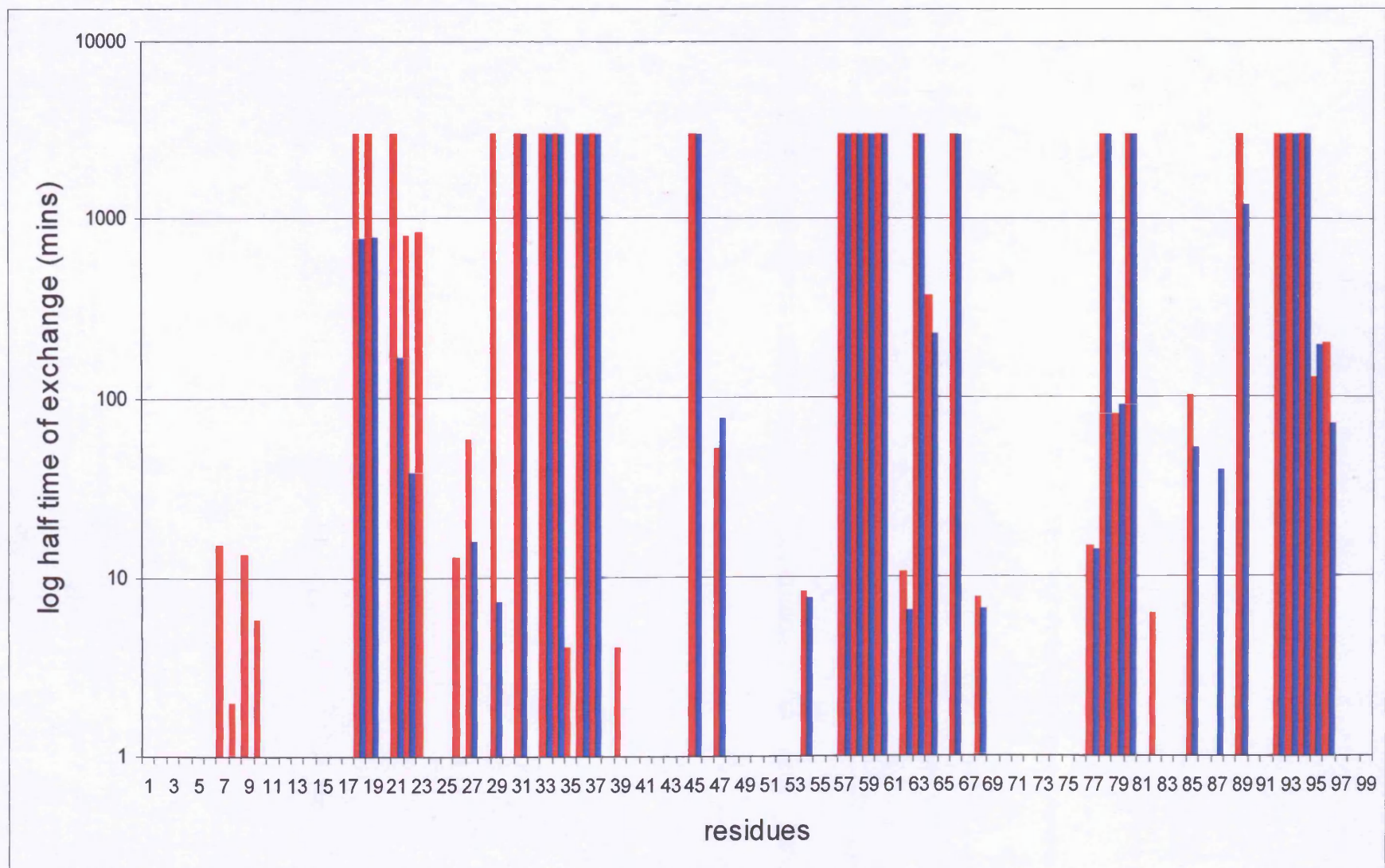
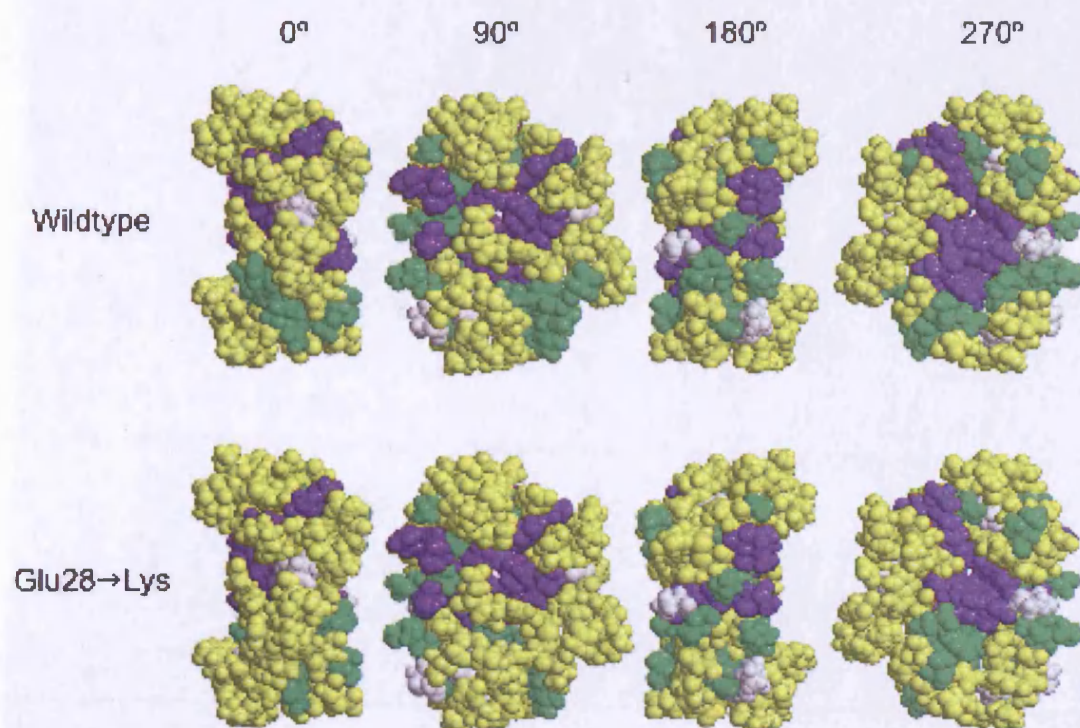


Figure 9.14: The residues protected from D₂O exchange mapped onto the best energy wildtype MSP1₁₉ NMR structure.

The residues where the amide proton was protected from exchange are mapped onto the best energy wildtype MSP1₁₉ NMR structure for the wildtype protein (top panel) and Glu28→Lys MSP1₁₉ variant (bottom panel). Residues where the amide proton exchanged in less than 5 minutes and were therefore not protected from exchange are shown in yellow. Residues where the amide proton took longer than 3000 minutes to exchange and were therefore protected from exchange are shown in purple. Residues where the amide proton took between 5 minutes and 3000 minutes to exchange and were therefore partially protected from exchange are shown in green. Proline residues are shown in white as they do not have an NH residue to exchange.





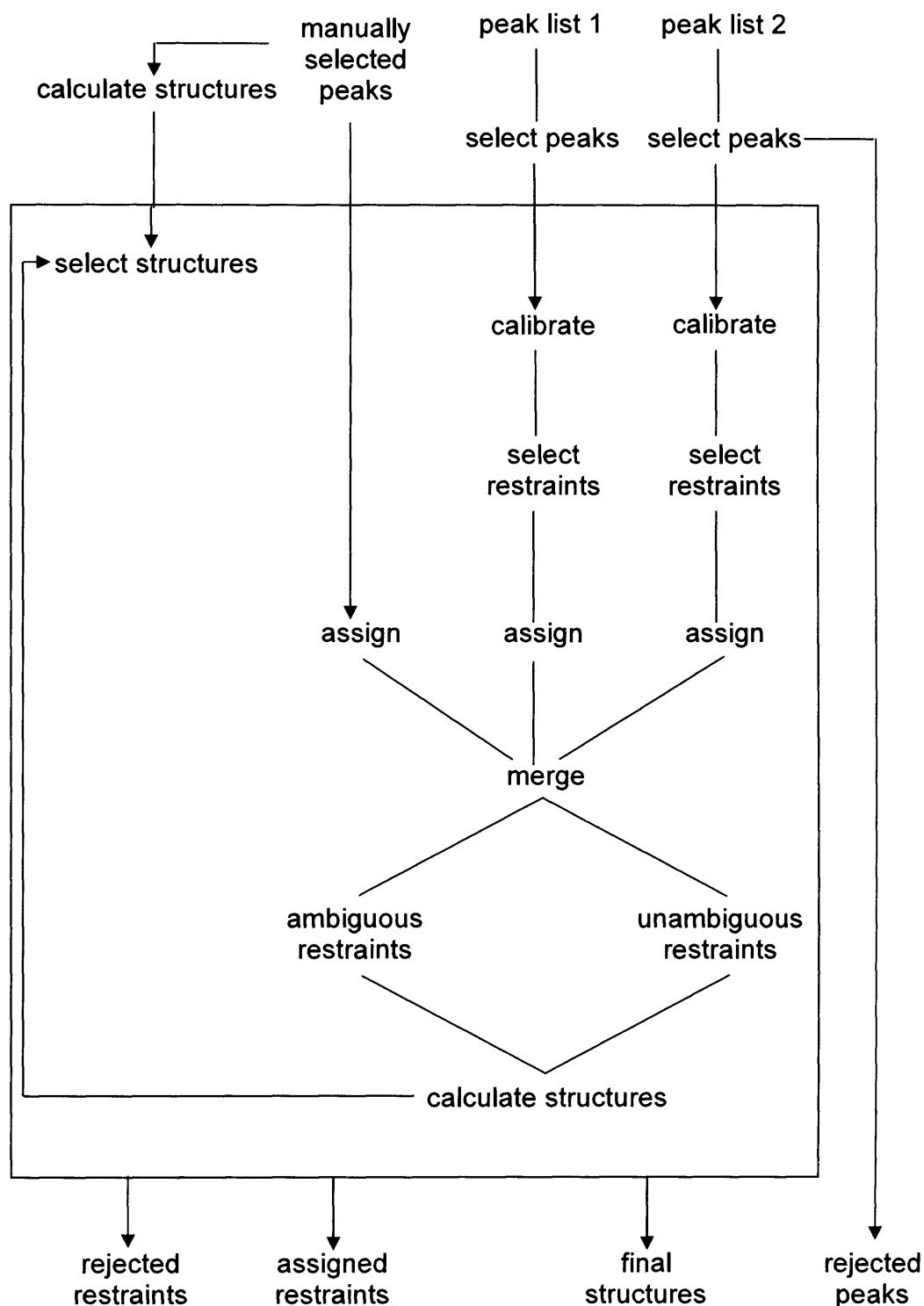
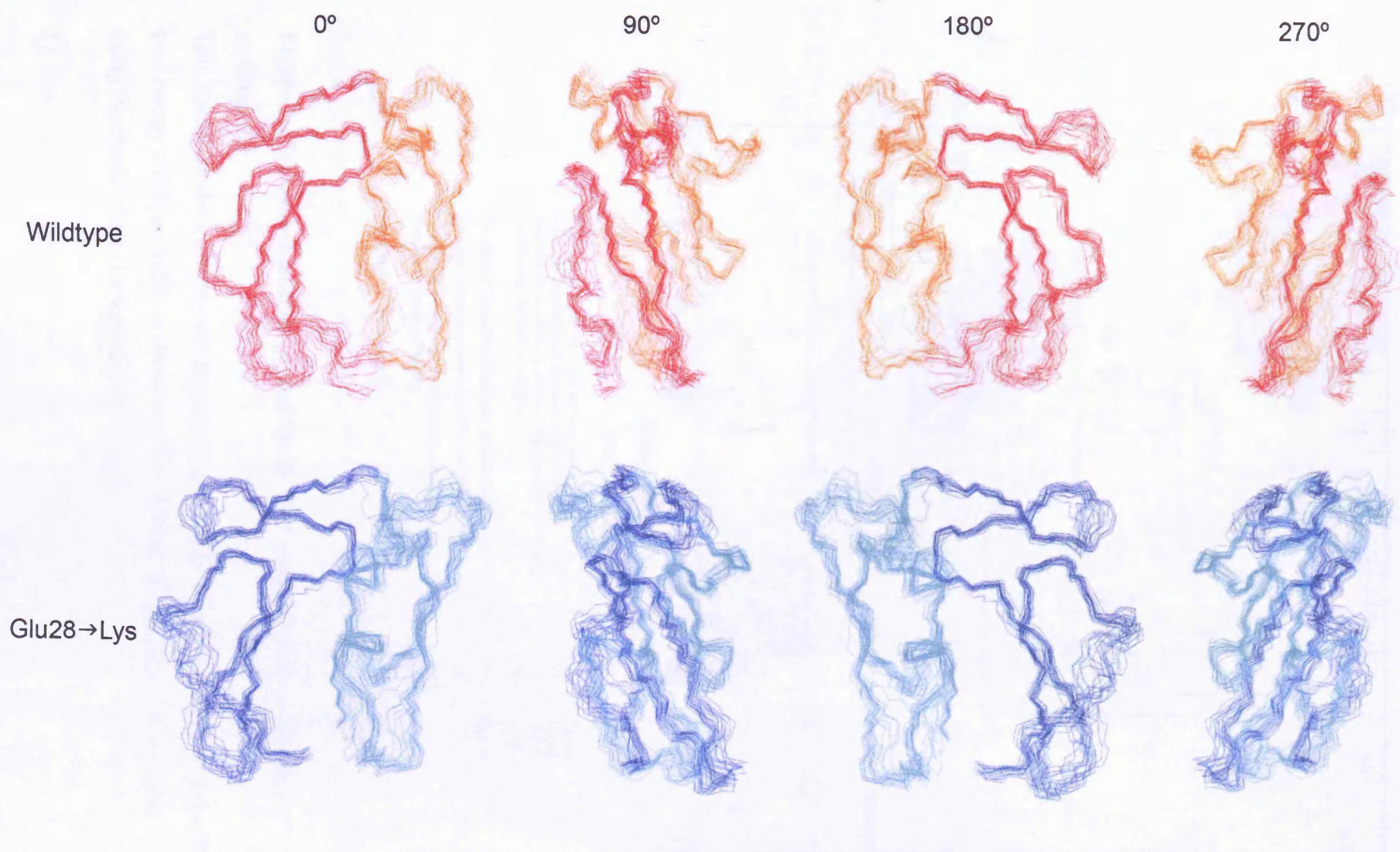


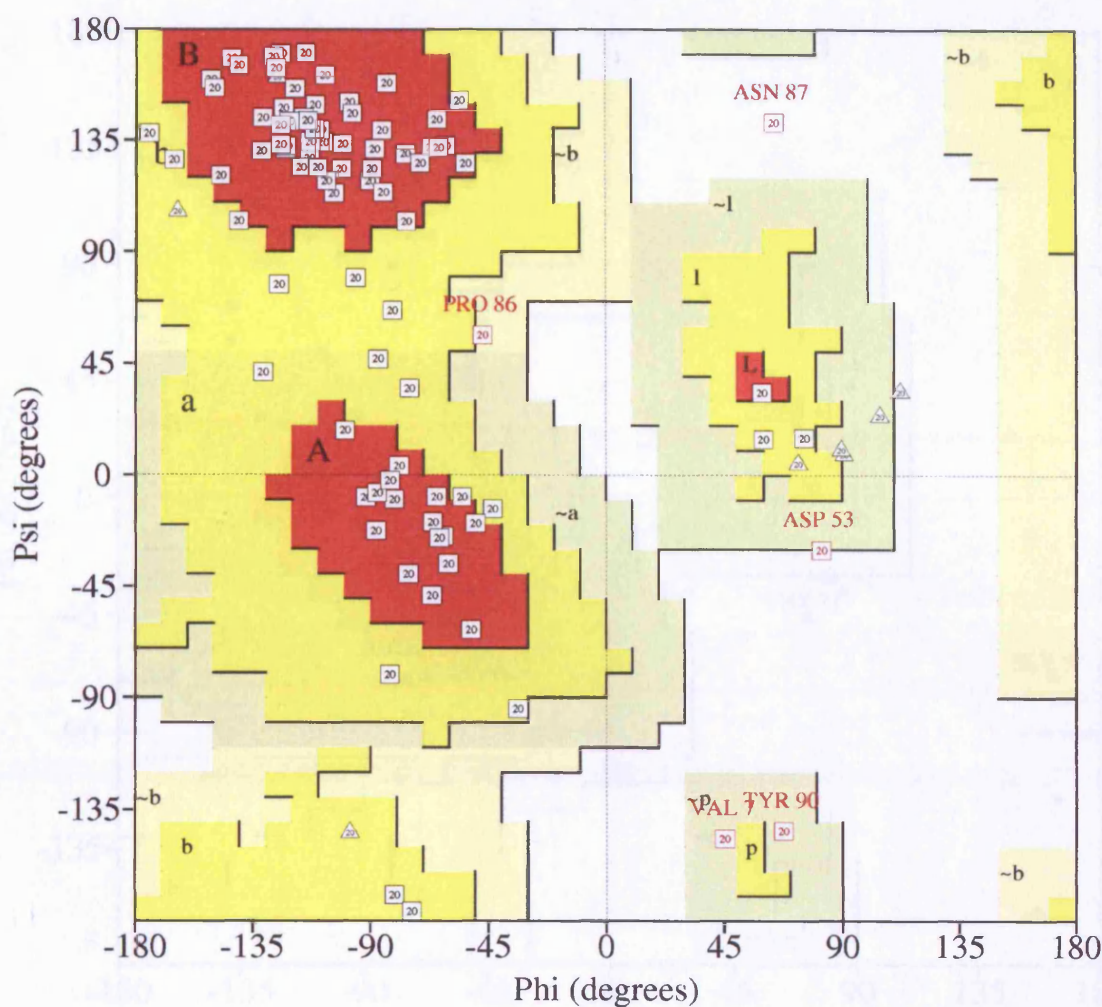
Figure 9.15: Schematic overview of the operation performed by ARIA.

This figure is based on figure 1 (Nilges *et al.*, 1997).

Figure 9.16: 20 best energy structures of wildtype and Glu28→Lys MSP1₁₉ variant.

The structures of wildtype and Glu28→Lys MSP1₁₉ variant were calculated using ARIA (Nilges, 1995, Nilges & O' Donoghue, 1998). The 20 best energy structures were displayed and superimposed using residues 10-90 using Insight II (Dayringer *et al.*, 1986). For wildtype MSP1₁₉ the first EGF domain is in red and the second EGF domain is in orange. For Glu28→Lys MSP1₁₉ variant the first EGF domain is in blue and the second EGF domain is in cyan.





Plot statistics

Residues in most favoured regions [A,B,L]	1280	77.1%
Residues in additional allowed regions [a,b,l,p]	300	18.1%
Residues in generously allowed regions [~a,~b,~l,~p]	40	2.4%
Residues in disallowed regions	40	2.4%

Number of non-glycine and non-proline residues	1660	100.0%
Number of end-residues (excl. Gly and Pro)	20	
Number of glycine residues (shown as triangles)	180	
Number of proline residues	120	

Total number of residues	1980	

Figure 9.17: Ramachandran plot of the 20 best energy wildtype MSP1₁₉ structures.

This Ramachandran plot shows the psi and phi angles for all the residues in the 20 best energy wildtype MSP1₁₉ structures. The Ramachandran plot was created using Procheck NMR (Laskowski et al., 1996)

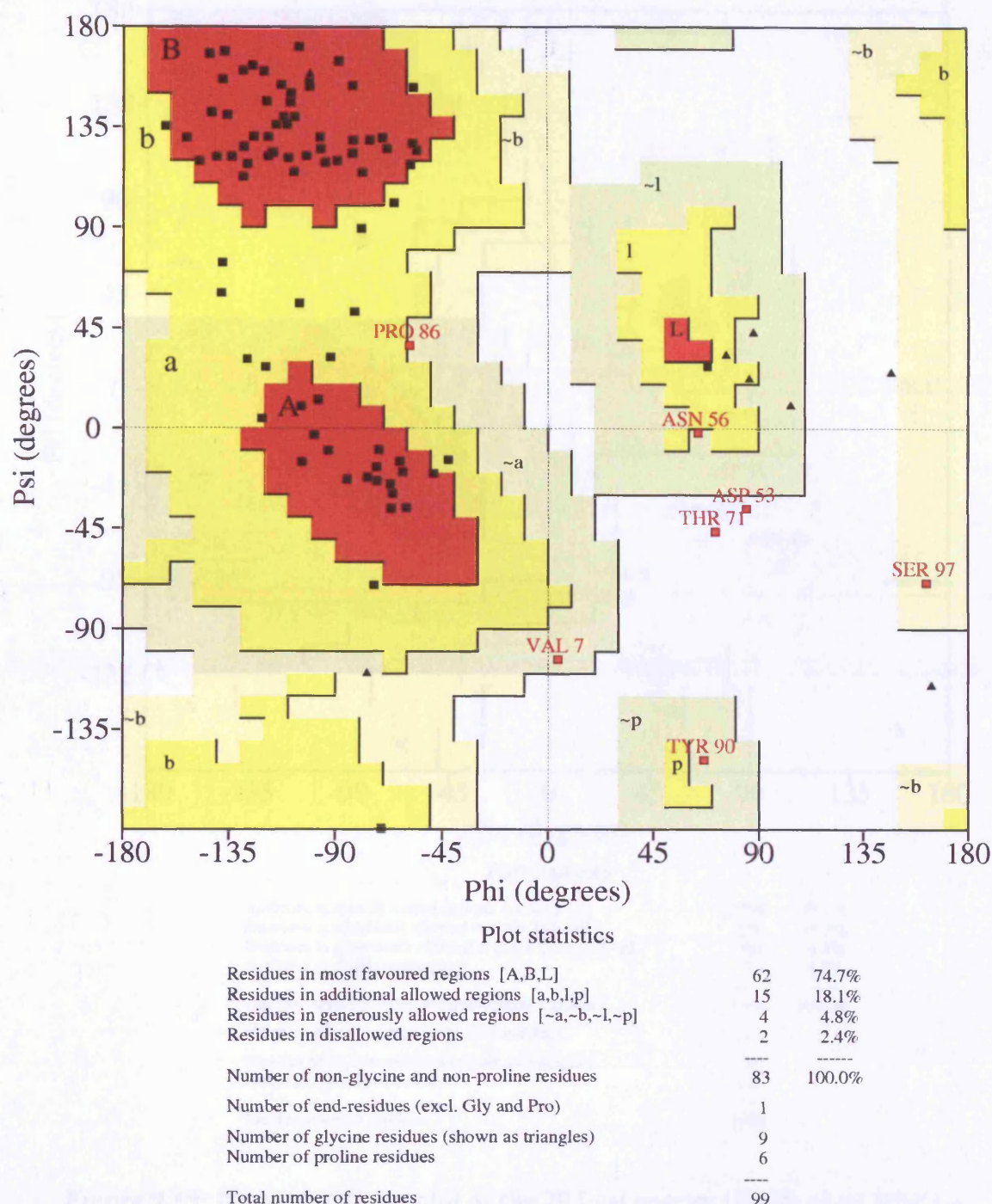


Figure 9.18: Ramachandran plot of the best energy wildtype MSP1₁₉ structure.

This Ramachandran plot shows the psi and phi angles for all the residues in the best energy wildtype MSP1₁₉ structure. The Ramachandran plot was created using Procheck NMR (Laskowski et al., 1996)

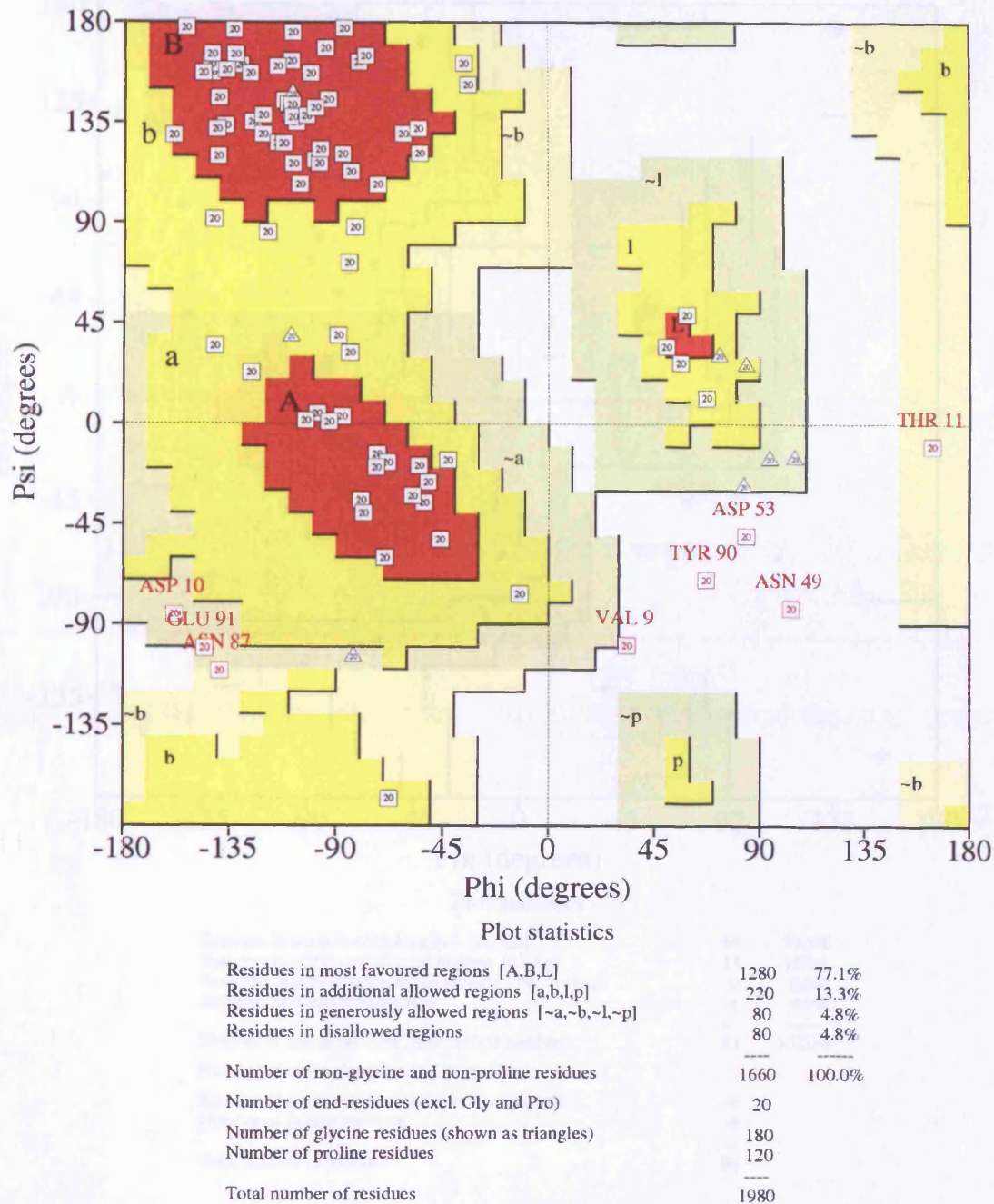


Figure 9.19: Ramachandran plot of the 20 best energy Glu28→Lys MSP1₁₉ variant structures.

This Ramachandran plot shows the psi and phi angles for all the residues in the 20 best energy Glu28→Lys MSP1₁₉ variant structures. The Ramachandran plot was created using Procheck NMR (Laskowski et al., 1996)

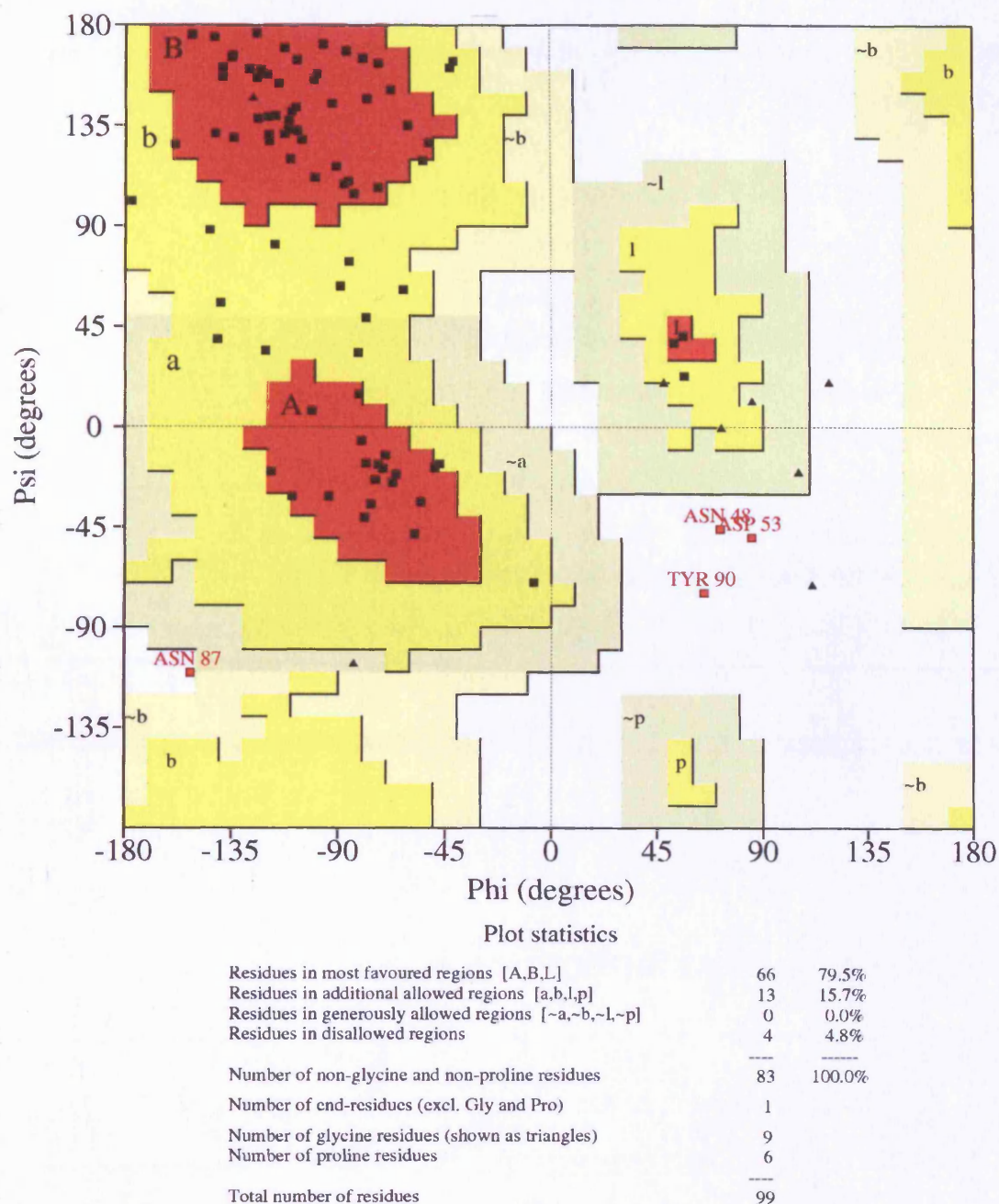


Figure 9.20: Ramachandran plot of the best energy Glu28→Lys MSP1₁₉ variant structure.

This Ramachandran plot shows the psi and phi angles for all the residues in the best energy Glu28→lys MSP1₁₉ variant structure. The Ramachandran plot was created using Procheck NMR (Laskowski et al., 1996)

Figure 9.21: Comparison of the best energy wildtype MSP1₁₉ NMR structure and the homology model.

A: The backbone structure of the best energy wildtype MSP1₁₉ NMR structure. The first EGF domain is shown in red and the second EGF domain is shown in orange.

B: The backbone structure of the homology model of MSP1₁₉. The first EGF domain is shown in pink and the second EGF domain is shown in purple.

C: The backbone structure of the homology model in pink superimposed onto the best energy wildtype MSP1₁₉ NMR structure in red using residues 10 to 90. The areas that are particularly different are highlighted. The areas include residues 8 to 11 shown in blue and cyan; residue 12 shown in bright yellow and pale yellow; residues 70 to 75 shown in black and grey and residues 80 to 88 shown in dark green and light green for the NMR structure and homology model respectively.

D: A comparison of the secondary structural elements of the best energy NMR structure in red and homology model in pink.

This figure was prepared using MolMol(Koradi *et al.*, 1996) and RasTop.

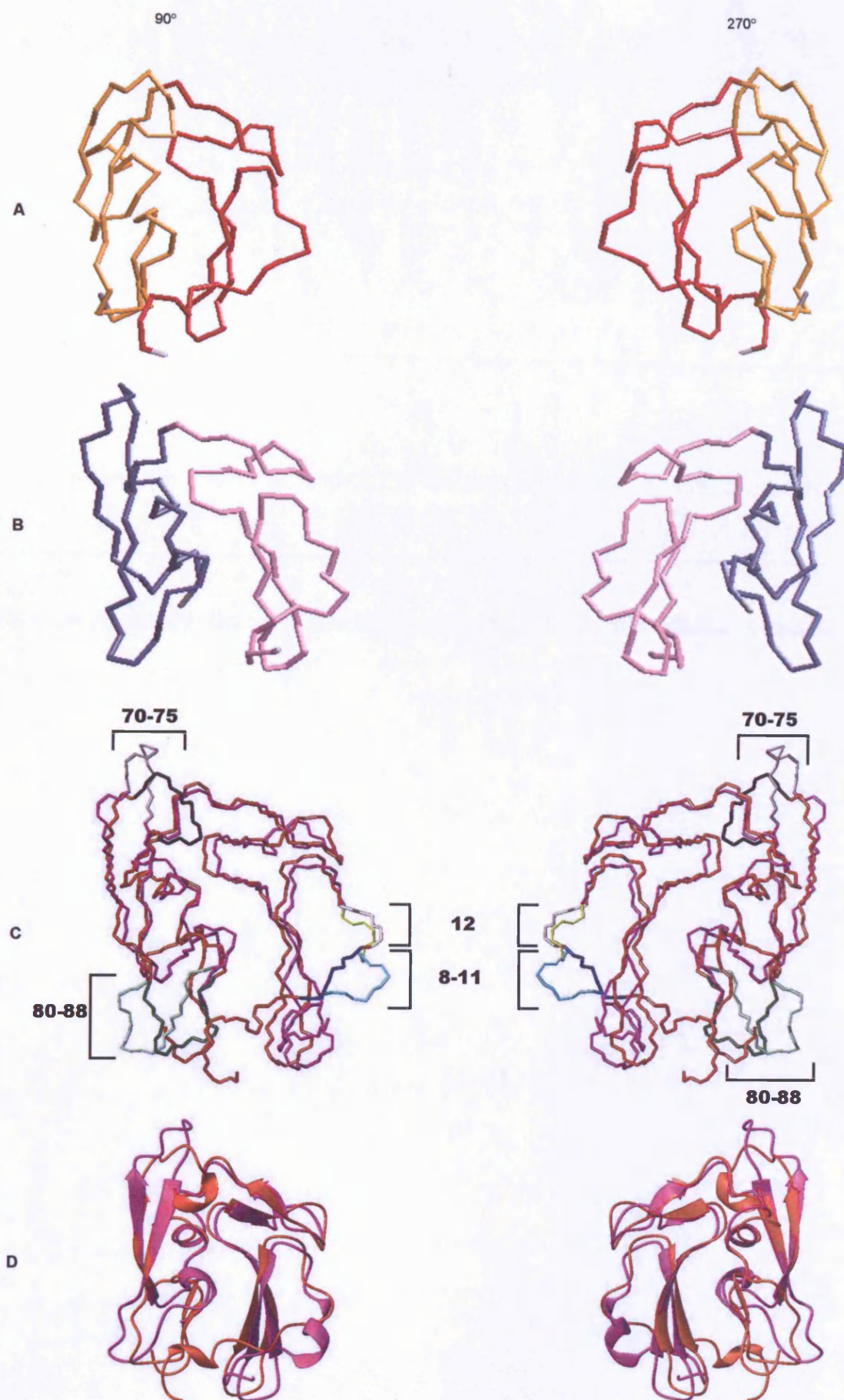


Figure 9.22: Comparison between the backbone structure of the first EGF domain of wildtype and Glu28→Lys MSP1₁₉.

The first EGF domain of wildtype in red is superimposed onto the Glu28→Lys MSP1₁₉ in blue using residue 8 to 48. This figure was prepared was prepared using Swiss Model (Guex & Peitsch, 1997).

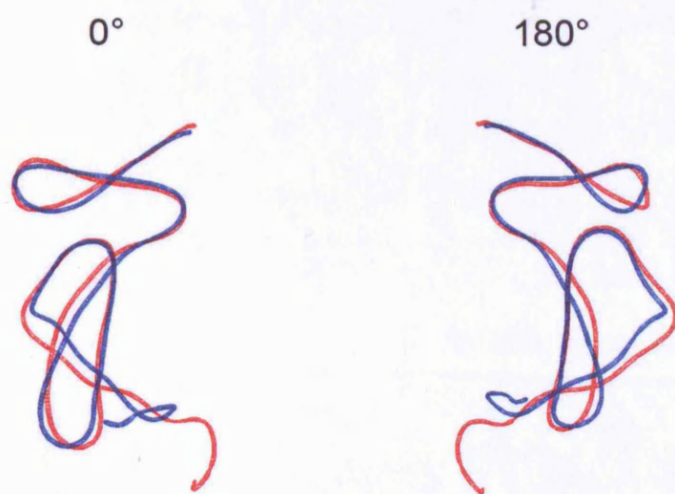


Figure 9.23: Comparison of the orientation of the residues in wildtype and Glu28→Lys MSP1₁₉ variant that have moved 0.2 ppm in the ¹⁵N-HSQC spectrum.

The side chains of the residues that have moved 0.2 ppm in the Glu28→Lys ¹⁵N-HSQC spectrum compared to the wildtype spectrum are shown on the best energy structures for wildtype and Glu28→Lys MSP1₁₉ variant to compare the orientation of the residues.

The first EGF domain of the wildtype structure is in red and the second EGF domain is in orange. The first EGF domain of Glu28→Lys MSP1₁₉ variant is in blue and the second EGF domain is in cyan. The residues are coloured according to their properties as follows: positively charged residues are in blue; negatively charged residues are in red; aromatic residues are in yellow and all other residues are in green.

This figure was prepared using RasTop.

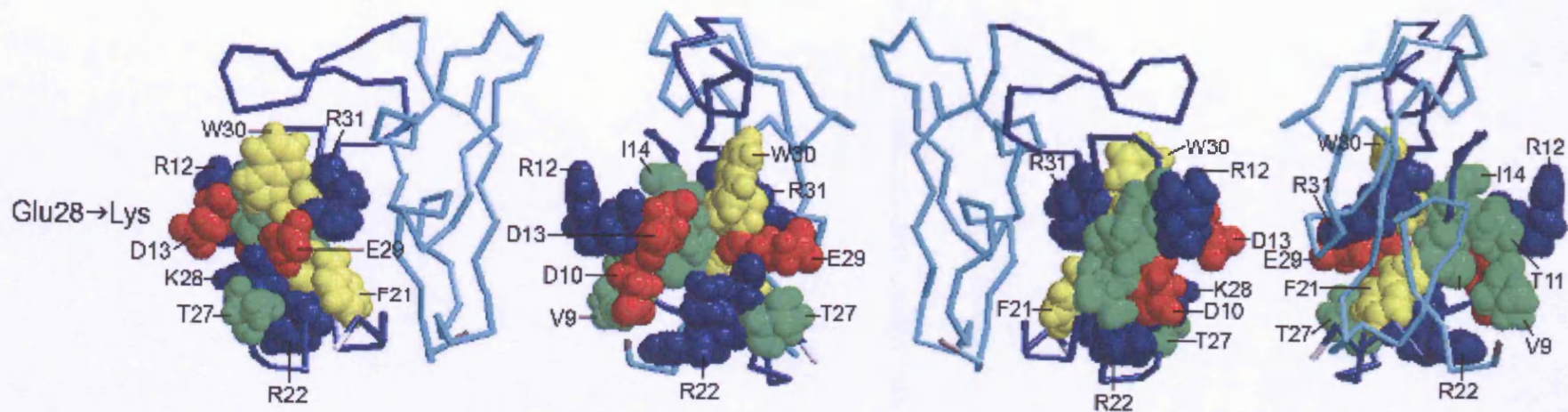
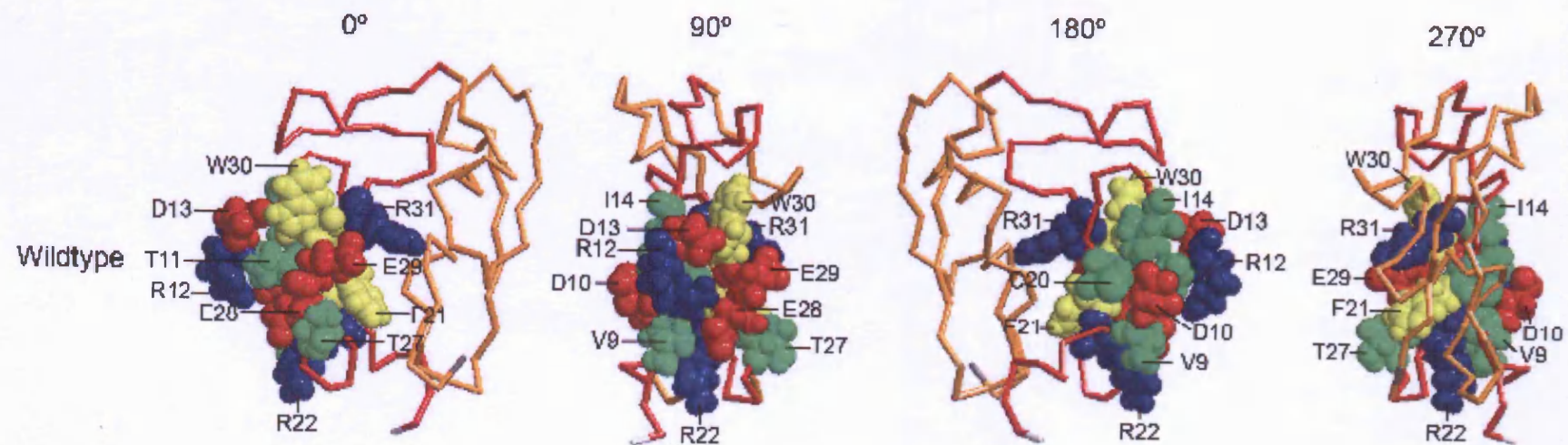


Figure 9.24: Comparison of the orientation of the charged residues in wildtype and Glu28→Lys MSP1₁₉ variant that have moved 0.2 ppm in the ¹⁵N-HSQC spectrum.

The side chains of the charged residues that have moved 0.2 ppm in the Glu28→Lys ¹⁵N-HSQC spectrum compared to the wildtype spectrum are shown on the best energy structures for wildtype and Glu28→Lys MSP1₁₉ variant to compare the orientation of the residues.

The first EGF domain of the wildtype structure is in red and the second EGF domain is in orange. The first EGF domain of Glu28→Lys MSP1₁₉ variant is in blue and the second EGF domain is in cyan. The residues are coloured according to their properties as follows: positively charged residues are in blue and negatively charged residues are in red.

This figure was prepared using RasTop.

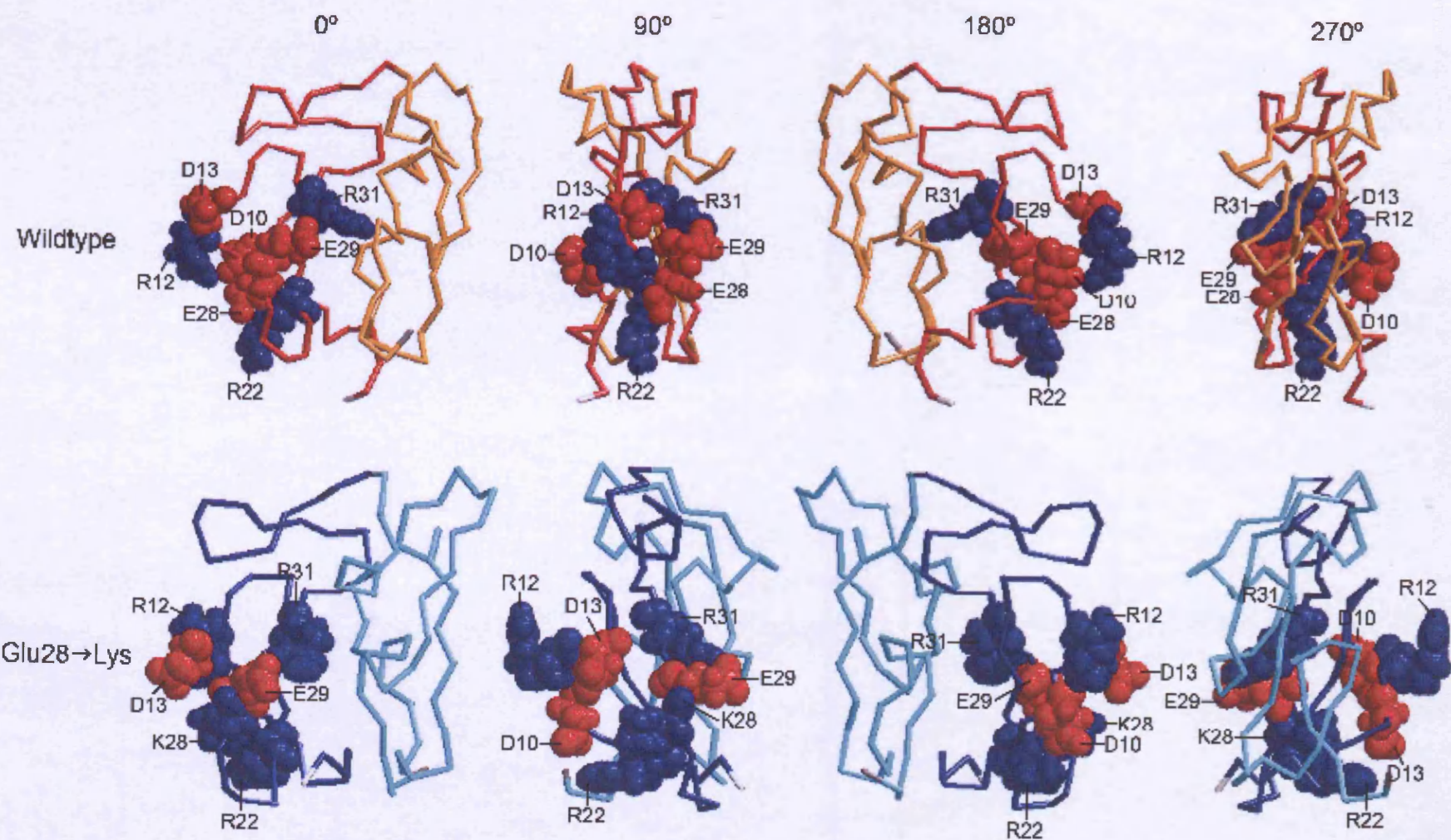
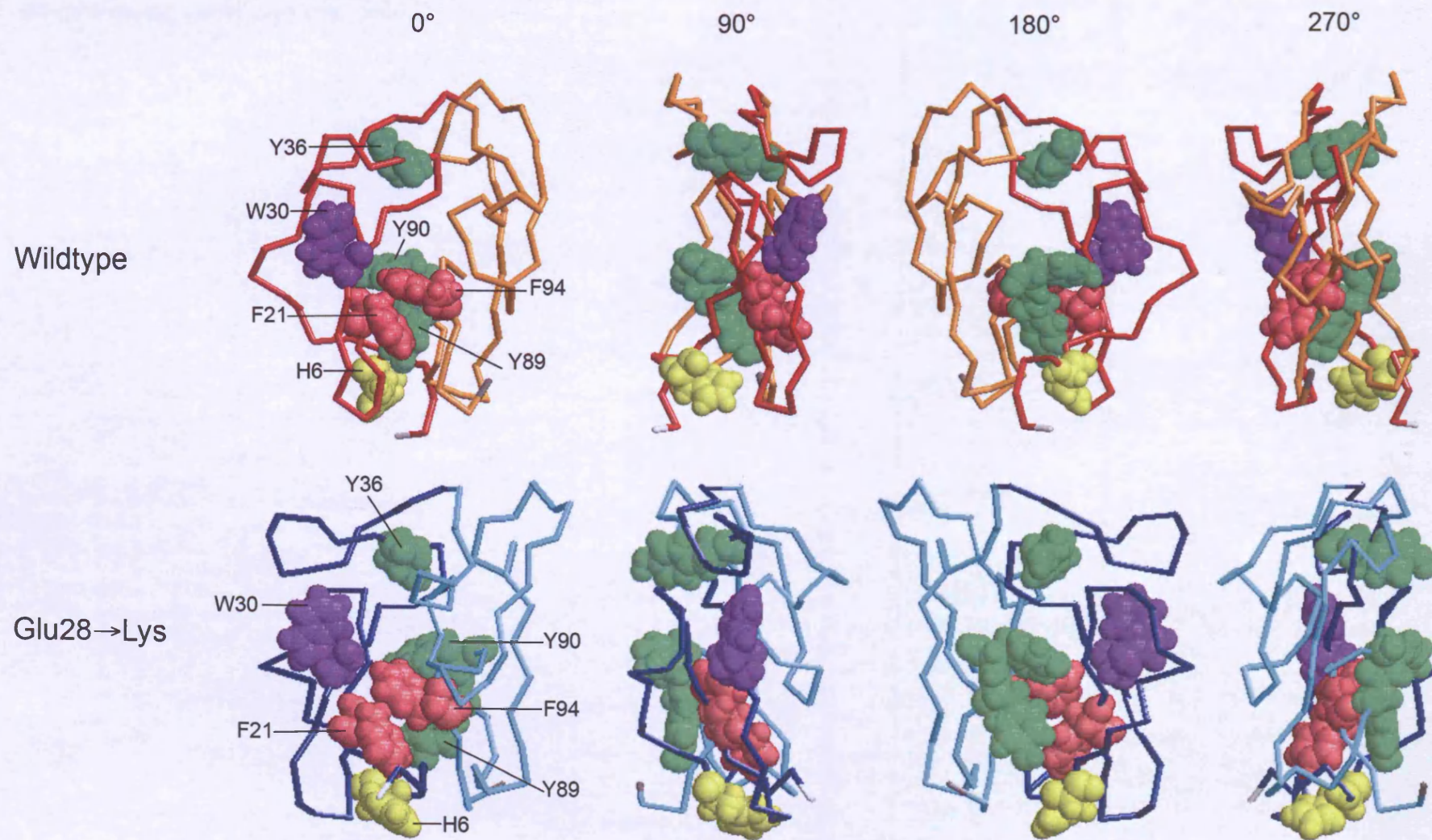


Figure 9.25: Comparison of the orientation of the aromatic residues in wildtype and Glu28→Lys MSP1₁₉ variant.

The side chains of the aromatic residues are shown on the best energy structures for wildtype and Glu28→Lys MSP1₁₉ variant to compare the orientation of the residues.

The first EGF domain of the wildtype structure is in red and the second EGF domain is in orange. The first EGF domain of Glu28→Lys MSP1₁₉ variant is in blue and the second EGF domain is in cyan. The residues are coloured according to the type of aromatic residue: phenylalanine residues are shown in pink; tyrosine residues are shown in green; tryptophan residues are shown in purple and histidine residues are shown in yellow.

This figure was prepared using RasTop.



9.7 Discussion

The quality of the NMR structures of the wildtype *P. yoelii* MSP1₁₉ and Glu28→Lys MSP1₁₉ variant that have been presented in this chapter are of comparable quality to the published NMR structures for *P. falciparum* and *P. vivax* MSP1₁₉. In the Ramachandran plots for the wildtype and Glu28→Lys MSP1₁₉ variant 77.1 % of residues have their phi and psi angles in the most favoured regions. The published NMR structures for *P. falciparum* (Morgan *et al.*, 1999) and *P. vivax* (Babon *et al.*, 2007) have 54.9 % and 69 % of their residues with phi and psi angles in the most favoured regions.

The comparison of the homology model with the wildtype *P. yoelii* MSP1₁₉ NMR structure has shown that the homology model can predict the overall fold of the protein but cannot accurately predict all areas of the protein structure and the orientation of the amino acid side chains. This may be because *P. falciparum*, *P. cynomolgi* and *P. knowlesi* MSP1₁₉ structures that were used to make the homology model all share the EGF motif consensus sequence but the rest of the areas of the sequence have lower sequence similarity. This would make it difficult for the Swiss Model Homology Modelling server (Guex & Peitsch, 1997) to accurately predict the orientation of the side chains. The differences between the homology model the NMR structure around residues 8 to 12 suggest why the *in silico* variation of residue 28 was unable to accurately predict the changes in this area that were seen in the Glu28→Lys MSP1₁₉ variant NMR structure.

The D₂O exchange studies have identified differences in the hydrogen bonding between the wildtype and Glu28→Lys MSP1₁₉ variant NMR structures. The rapid exchange rates from residues 7 to 10 seen in the Glu28→Lys MSP1₁₉ variant suggests these residues are not involved in hydrogen bonding but in the wildtype there is slower exchange indicating protection due to the presence of hydrogen bonds in this area. This suggests that in this area of the structure the hydrogen bonds have broken in Glu28→Lys MSP1₁₉ variant. This may mean that the structure in this area for Glu28→Lys MSP1₁₉ variant is not as rigid as the structure in the wildtype. This

difference in hydrogen bonding is in agreement with data for the ¹⁵N-HSQC NMR spectra which suggested that the hydrogen bonds in residues 9 and 10 had been broken because they had shifted upfield between the wildtype and Glu28→Lys MSP1₁₉ variant. In the wildtype ¹⁵N-HSQC they were in a position where they were shifted downfield in comparison to the other residues in the protein indicating that they were involved in hydrogen bonding because of the shifting of the residues, whereas in Glu28→Lys MSP1₁₉ variant they were further upfield suggesting they were not involved in a hydrogen bond. The area between residues 21 and 29 also indicated a difference in protection in the D₂O exchange experiments. This difference in protection in this area could suggest the variation to residue 28 has opened up this area of the protein and made it less rigid resulting in making the area more accessible to the solvent.

The comparison of the wildtype and Glu28→Lys MSP1₁₉ variant NMR structures show clear differences in the first EGF domain. The data suggests glutamic acid 28 is in a negatively charged area of the protein and by replacing the negatively charged glutamic acid 28 with a positively charged lysine this has had an affect on the surrounding charged residues. The data suggests that residue 28 is interacting with arginine 12 in the wildtype structure and that when residue 28 is changed to a lysine this then repels arginine 12. The movement of arginine 12 influences the nearby negatively charged aspartic acid residues 10 and 13 with residue 13 appearing to swap places with arginine 12. The movement of aspartic acid 10 and 13 may be a result of an attraction from the positively charged lysine 28 side chain whereas the glutamic acid 28 in the wildtype would have repelled the aspartic acid residues. The data could suggest that the area around glutamic acid 28 is forming a charged antibody binding pocket and that antibodies could be recognising this area due to the interaction with the charged side chains. The wider implication of the variation to residue 28 altering the orientation of the two EGF domains relative to one another could be an affect on the orientation of the aromatic amino acids in the interface between the two domains. The aromatic residues phenylalanine 21 and 94 are in the interface between the two domains and are close to the area that has been disrupted by the residue 28 variation. The side chain orientation of these two residues has altered between the wildtype and Glu28→Lys MSP1₁₉ variant NMR structures. This could be contributing to altering the orientation of the two EGF domains because phenylalanine is a large bulky residue. The movement of the aromatic

residues could explain why such a large number of residues had shifted in the 2D ¹⁵N-HSQC spectrum for Glu28→Lys MSP1₁₉ variant. The ring current effect of the aromatic ring of the aromatic residue affects amino acids that are in the same plane as the aromatic ring. This would mean that where the orientation of the aromatic residue has altered this could change which of the surrounding amino acid chemical shifts are influenced by the ring current effect.

The NMR structures presented in this chapter have shown that the alteration of residue 28 from a glutamic acid to a lysine has had a significant affect on the first EGF domain of the protein concentrated around the charged residues around residue 28. This suggests that residue 28 has a vital role in shaping the structure of the protein in that area. It indicates that the affects on antibody binding of the Glu28→Lys MSP1₁₉ variant cannot be a direct result of binding to residue 28 itself but due to the alterations of the structure of the protein and the charge distribution in the area of the protein around residue 28. The implications of the changes in the charge distribution for the immunology of the protein will be discussed in detail in chapter 10.

Chapter 10: Discussion

10.1 Introduction

In this thesis, I have mapped the antibody binding sites of antibodies to *P. yoelii* MSP1₁₉ by investigating the affect of single and double amino acid changes to the protein on antibody binding. The results of the antibody binding studies by western blotting, ELISA and surface plasmon resonance that were obtained are summarised in table 10.1.

Table 10.1: Summary of the affect of the amino acid variations to MSP1₁₉ on monoclonal antibody binding

Variant	B6	F5	B10
Arg12→Leu	++	+	++
Lys16→Glu	-	-	++
Asn17→His	+	++	++
Glu28→Lys	+	-	+
double Lys16→Glu/Glu28→Lys	-	-	+
Glu28→Gln	+	-	+
- → variation abolishes antibody binding + → variation partially reduces antibody binding ++ → variation has no affect on antibody binding			

The MSP1₁₉ variant proteins were used in *in vivo* immunisation studies to investigate the affect of the variation on the ability of the proteins to protect against challenge infection and to see if this correlated with the *in vitro* data. The *in vivo* immunisation studies showed that the Arg 12→Leu, Lys16→Glu and Asn17→His MSP1₁₉ variant proteins protected against subsequent parasite challenge in the same way as the wildtype protein. This showed that even though residue 16 abolished antibody binding *in vitro* to B6 and F5 antibody it did not translate to an affect *in vivo*. Immunisation with

Glu28→Lys, double Lys16→Glu/Glu28→Lys and Glu28→Gln MSP1₁₉ variant proteins did not protect against parasite challenge *in vivo*. ELISA experiments examining the antibody titres of the mice immunised with the wildtype and MSP1₁₉ variants showed that there was no significant difference between the antibody titres for the mice immunised with wildtype MSP1₁₉ and the MSP1₁₉ variants. This meant that the overall antibody titre could not be used to predict the ability of the protein to protect against parasite challenge. When the ELISA titres for the mice immunised with Glu28→Lys, double Lys16→Glu/Glu28→Lys and Glu28→Gln MSP1₁₉ variant against wildtype protein were compared to those against the variant protein it was found that the antibody titres to the variant protein were slightly higher. This suggested that a small proportion of the antibody response to MSP1₁₉ was specific to the area of the protein altered by the residue 28 variants. The data suggested that the fine specificity of the antibody response is more important than the overall antibody titre in determining whether the antibody response will be able to protect against parasite challenge.

Structural studies were carried out to compare the structure of the wildtype and variant proteins. This confirmed that for the residue 12, 16 and 17 variants there were no significant structural differences and the results of the *in vitro* and *in vivo* experiments were a direct result of the changes to the individual residue. The structural studies for the residue 28 variants suggested there was significant structural perturbation as a result of these variations. The structure of wildtype and Glu28→Lys MSP1₁₉ variant confirmed that the residue 28 variation had significantly altered the first EGF domain and the residues in the interface between the two EGF domains. This suggests that residue 28 may have a vital structural role.

In this chapter, I will discuss the structure-function relationship between wildtype MSP1₁₉ and Glu28→Lys MSP1₁₉ variant. The structure of the *P. yoelii* MSP1₁₉ will be compared to the structures of MSP1₁₉ from other species. The amino acid variations will be compared to those in the literature for other species to examine whether similar residues are involved. I will also discuss the implications of the work presented in this thesis for other studies and future work.

10.2 Electrostatic potential of wildtype and Glu28→Lys MSP1₁₉ variant

The electrostatic potentials for the wildtype and Glu28→Lys MSP1₁₉ variant proteins were determined using MOLMOL (Koradi *et al.*, 1996) and mapped onto the surface of the proteins. The electrostatic surface potential is shown in figure 10.1 for wildtype and Glu28→Lys MSP1₁₉ variant. This shows that there are distinct clusters of positively and negatively charged residues on the surface of the two proteins. For the wildtype protein in the 180° rotation there is a large positively charged area at the top of the protein with a large negatively charged area below including glutamic acid 28 and aspartic acid 24. For the Glu28→Lys MSP1₁₉ variant there is a clear difference in the charge distribution on the surface compared to the wildtype protein especially around residue 28. In the Glu28→Lys MSP1₁₉ variant the positively charged lysine 28 has introduced a positively charged area where there was just a negatively charged area in the wildtype protein. This change is seen clearly in the 90° rotation where lysine 28 in the Glu28→Lys MSP1₁₉ variant is in a positively charged area with a few negative residues around it and in the wildtype protein glutamic acid 28 is in a negatively charged area with a positively charged area next to it including arginine 12.

10.3 Comparison of electrostatic potential of wildtype MSP1₁₉ from different species

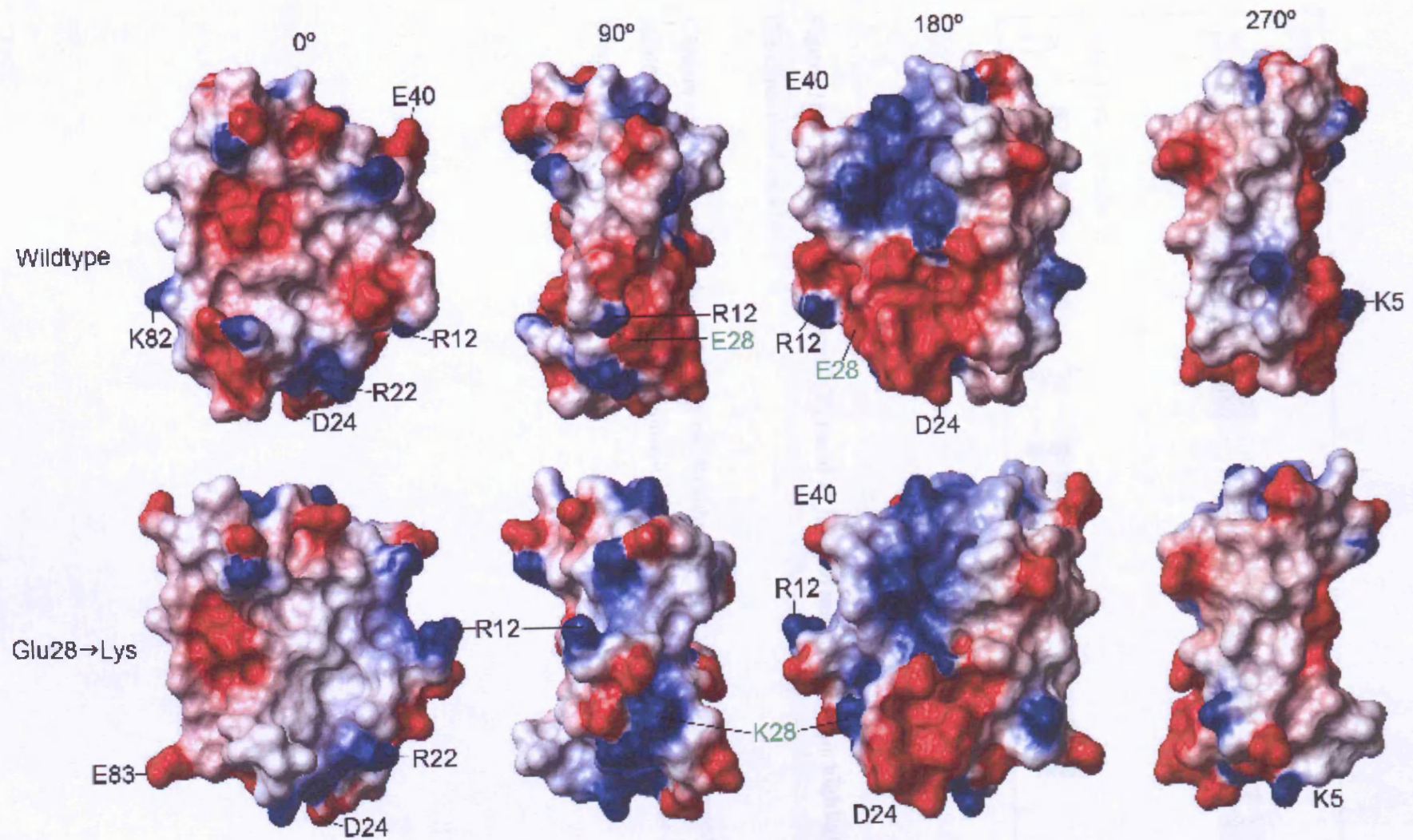
Figure 10.2 shows an alignment of the amino acid sequence of wildtype *P. yoelii* MSP1₁₉ and wildtype *P. falciparum* MSP1₁₉. This shows that the primary EGF structural motif (highlighted in green) is conserved across the species. The backbone structure of *P. yoelii* MSP1₁₉ and *P. falciparum* MSP1₁₉ was superimposed using residues 10 to 90 in MOLMOL (Koradi *et al.*, 1996) to compare similarities between the two proteins (shown in figure 10.3). The backbone for *P. yoelii* MSP1₁₉ is shown in red and the backbone for *P. falciparum* MSP1₁₉ is shown in green. This shows that the overall structure of the two proteins is conserved and the overall secondary structural elements are the same for the two proteins. The biggest differences between the two proteins occur in the looped regions where there are no secondary structural elements.

The electrostatic potentials for *P. yoelii* MSP1₁₉ and *P. falciparum* MSP1₁₉ proteins was determined using MOLMOL(Koradi *et al.*, 1996) and mapped onto the surface of the proteins (shown in figure 10.4). This shows that there are distinct clusters of positively and negatively charged residues on the surface of the two proteins but the distribution of charges is different. The 0° face is very similar for the two proteins and on the 90° face the *P. yoelii* protein has more charged residues particularly a cluster of negatively charged residues around glutamic acid 28. The 180° face shows more differences between the two proteins. The *P. yoelii* protein has very distinct positively and negatively charged areas with a positively charged area at the top and a negatively charged area below. The *P. falciparum* protein has a negatively charged cluster of residues but has individual positively charged residues instead of a distinct cluster of positively charged residues.

In the literature the electrostatic potential of *P. vivax*, *P. cynomolgi* and *P. knowlesi* has been compared to *P. falciparum* MSP1₁₉ (Babon *et al.*, 2007, Garman *et al.*, 2003, Pizarro *et al.*, 2003). The electrostatic potential for *P. yoelii* MSP1₁₉ was also compared to the electrostatic potential of *P. vivax*, *P. cynomolgi* and *P. knowlesi* MSP1₁₉ (as shown figure 10.5) to see if the electrostatic potential was more similar to these proteins. This figure shows that the proteins have a different distribution of charged residues on the surface compared to *P. yoelii* and each other. *P. vivax*, *P. cynomolgi* and *P. knowlesi* MSP1₁₉ all have large areas of negatively charged residues and individual or smaller groups of positively charged residues but the distribution of these areas differs between the proteins. For *P. yoelii* a large area of positive charge at the top of 180° face is only seen for this species.

Figure 10.1: Comparison of electrostatic potential of wildtype and Glu28→Lys MSP1₁₉ variant

The electrostatic potentials of wildtype and Glu28→Lys MSP1₁₉ variant are mapped onto the molecular surface of the proteins. The electrostatic potentials were calculated using MOLMOL (Koradi *et al.*, 1996). Red represents negative charge and blue represents positive charge.



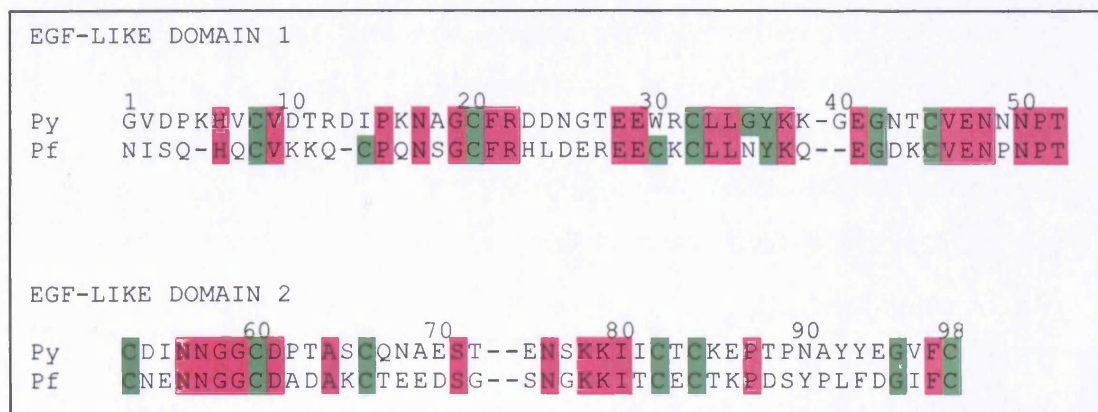


Figure 10.2: Alignment of *P. yoelii* (Py) and *P. falciparum* (Pf) MSP1₁₉ highlighting the conserved residues.

Conserved residues are highlighted in pink. Residues that are part of the conserved EGF motif are highlighted in green. The alignment is based on figure 3 of Benjamin *et al.* (Benjamin *et al.*, 1999)

Figure 10.3 Comparison of *P. yoelii* and *P. falciparum* MSP1₁₉ NMR structures.

The backbone of the best energy NMR structure of *P. yoelii* MSP1₁₉ (shown in red) superimposed on top of the backbone of the best energy NMR structure of *P. falciparum* MSP1₁₉ (Morgan *et al.*, 1999) (shown in green). The backbones are superimposed using residues 10–90 in MOLMOL (Koradi *et al.*, 1996) and the secondary structural elements are displayed.

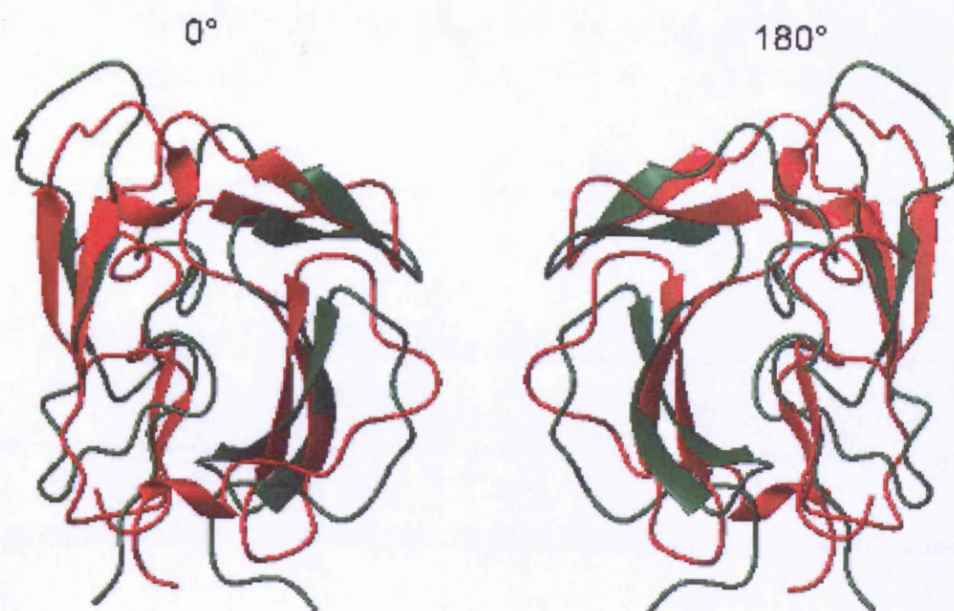


Figure 10.4: Comparison of electrostatic potential of *P. yoelii* and *P. falciparum* MSP1₁₉

The electrostatic potentials of *P. yoelii* and *P. falciparum* MSP1₁₉ are mapped onto the molecular surface of the proteins. The electrostatic potentials were calculated using MOLMOL (Koradi *et al.*, 1996). Red represents negative charge and blue represents positive charge. E37 labelled on the *P. falciparum* MSP1₁₉ structure is the equivalent of E40 in the *P. yoelii* MSP1₁₉ structure.

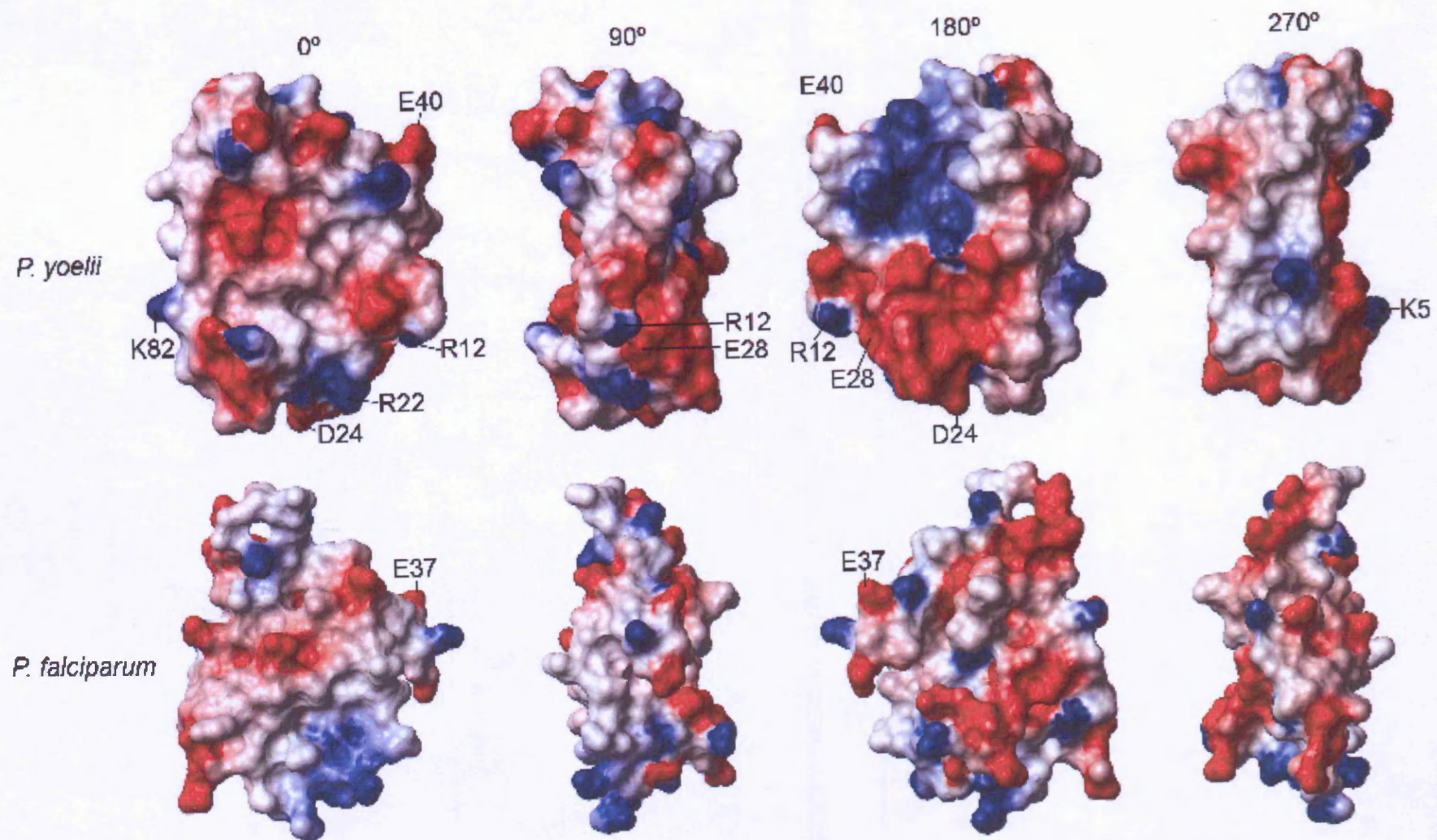
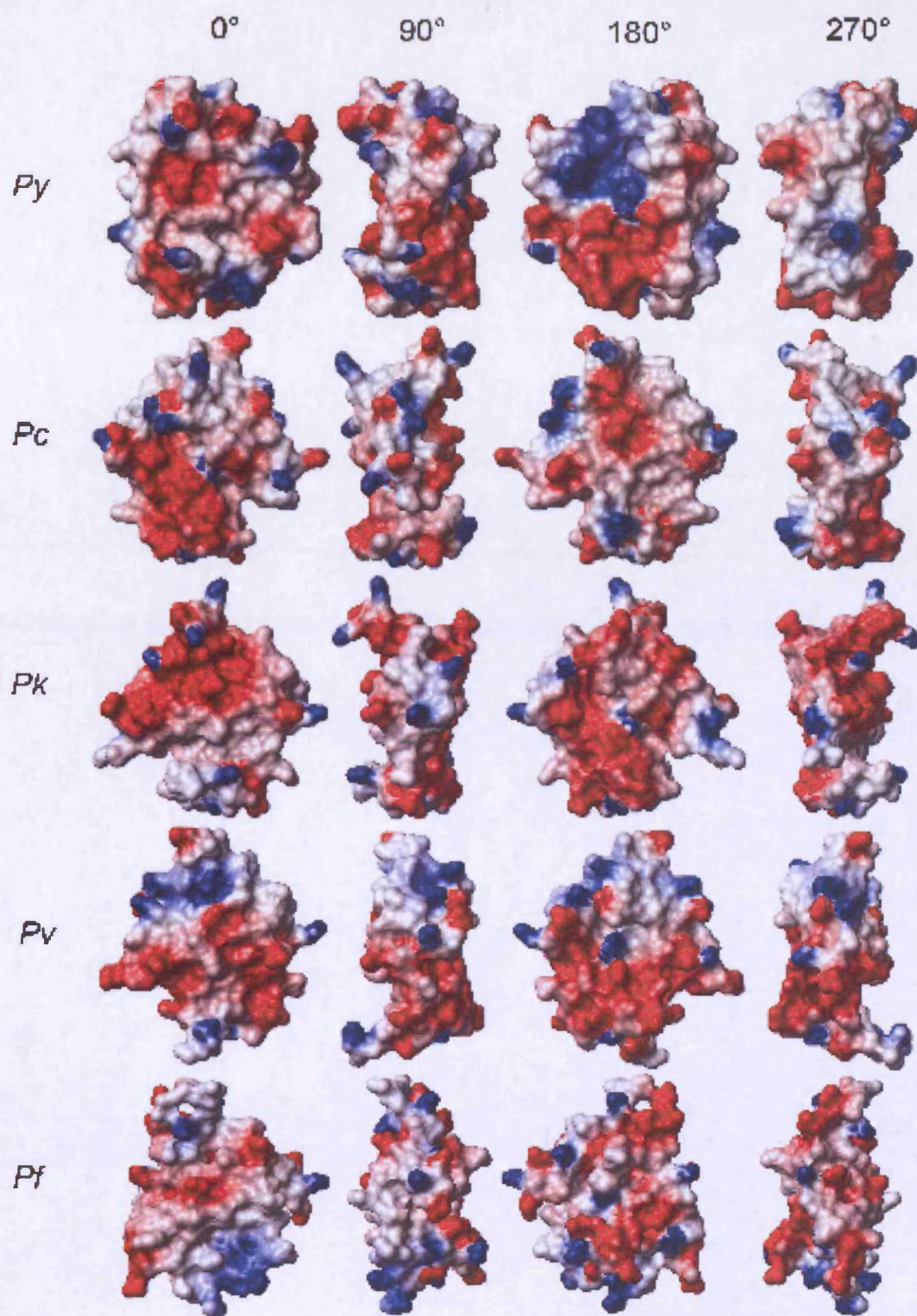


Figure 10.5: Comparison of electrostatic potential of MSP1₁₉ from different species.

The electrostatic potentials of *P. yoelii* (*Py*), *P. cynomolgi* (*Pc*) (Chitarra *et al.*, 1999), *P. vivax* (*Pv*) (Babon *et al.*, 2007), *P. knowlesi* (*Pk*) (Garman *et al.*, 2003) and *P. falciparum* (*Pf*) (Morgan *et al.*, 1999) MSP1₁₉ are mapped onto the molecular surface of the proteins. The electrostatic potentials were calculated using MOLMOL (Koradi *et al.*, 1996). Red represents negative charge and blue represents positive charge.



10.4 Discussion

*10.4.1 Structure function relationship of wildtype *P. yoelii* MSP1₁₉ and Glu28→Lys MSP1₁₉ variant*

The data have shown that by altering residue 28 to a lysine this has disrupted the electrostatic potential of the protein in the area around residue 28. These data are in agreement with the structural NMR studies presented in chapter 9, which show that there had been significant movement of the charged residues in the first EGF domain. The electrostatic potential data explain why changing from a negatively charged glutamic acid residue to a positively charged lysine residue can have such a significant impact on the protein. The 90° face of the protein (shown in figure 10.1) has a very different charge distribution in the Glu28→Lys variant and the wildtype protein; this would present a very different face to the antibodies and could explain why binding to B6, F5 and B10 was affected by this variation. The large number of charged residues on the surface of the protein in this area suggests that there is a charge interaction with antibodies binding to the surface. This would mean that if the recognition of the antibody binding site by the antibody relied on a charge interaction with the surface of the protein, the antibody would not recognise the Glu28→Lys protein as the same protein because it has a positively charged area where the wildtype protein had a negatively charged area. The positively charged area of the Glu28→Lys variant could potentially repel the antibody. The differences in electrostatic potential between the two proteins could also explain the differences in antibody titre to wildtype protein and Glu28→Lys MSP1₁₉ for mice immunised with Glu28→Lys MSP1₁₉. This is because antibodies that had been made to the area around residue 28 for the Glu28→Lys MSP1₁₉ protein would recognise a positively charged protein surface and therefore may be repelled by the negatively charged surface of the wildtype protein. Other areas of the wildtype and Glu28→Lys MSP1₁₉ have a very similar electrostatic potential and structure (as shown in the NMR studies in chapter 9), which would suggest that antibodies made to the rest of the Glu28→Lys MSP1₁₉ protein would be able to recognise and cross-react with wildtype MSP1₁₉.

The electrostatic potential for the wildtype protein could help to explain why the residue 12, 16 and 17 variants did not show large structural differences and did not affect *in vivo* protection. The Arg12→Leu variant involved changing an amino acid that is very close to residue 28 but it only involved losing a charge and is not found in the middle of the positively charged area. The Lys16→Glu variant involved changing from a positive to a negative charge but this residue is found in an area of the protein where there are few charged residues. This would mean that by changing the charge of this one residue it would be unlikely to have a huge affect on the neighbouring uncharged residues therefore keeping the structure around it intact. The significant affect that the Lys16→Glu variation has on B6 and F5 antibody binding could involve a specific charge interaction between the antibody and lysine that would not occur with the negatively charged glutamic acid residue. This could also explain why the residue 16 variant did not have an affect *in vivo* because antibodies made to the area of the protein could potentially cross react with the wildtype protein because the change is very localised.

10.4.2 Structural comparison between P. falciparum and P. yoelii

The structural comparison between *P. yoelii* MSP1₁₉ and *P. falciparum* MSP1₁₉ has shown that the overall secondary structural elements are conserved across the species. This would agree with the alignment of the amino acid sequences of *P. yoelii* and *P. falciparum* MSP1₁₉ which shows that the EGF structural motif is conserved across the species and the cysteine residues for five out of the six disulphide bonds in *P. falciparum* are conserved in *P. yoelii*. The differences between the loop regions of the proteins could be explained by the differences in sequences of the two proteins as the size and charges of the individual amino acids could shape these regions that are not held in a defined secondary structure. The electrostatic potential of the *P. yoelii* and *P. falciparum* MSP1₁₉ protein has shown that although there is overall conservation of the backbone structure of the two proteins there is not a conserved pattern of charges on the surface of the protein. The differences between the electrostatic potential of the two proteins could be explained by the difference in charges of the individual residues that make up the proteins.

The differences between the electrostatic potential of *P. yoelii* MSP1₁₉ and *P. vivax*, *P. cynomolgi* and *P. knowlesi* MSP1₁₉ suggest that the charge distribution on the surface of proteins is not conserved and is individual for each protein. This is in agreement with the published comparisons between *P. falciparum* and *P. vivax*, *P. cynomolgi* and *P. knowlesi* (Babon *et al.*, 2007, Garman *et al.*, 2003, Pizarro *et al.*, 2003), which showed that the electrostatic potentials of the proteins were different and that this may have been due to the low sequence similarity between the proteins.

10.4.3 Comparison of *P. yoelii* and *P. falciparum* antibody binding sites

Studies have been carried out by Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) and Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) to map the binding sites of *P. falciparum* MSP1₁₉ inhibitory antibodies. These data can be compared to the *P. yoelii* data presented in this thesis to determine whether there is a common area for inhibitory antibody binding. Figure 10.6 highlights the residues that affect inhibitory antibody binding in *P. falciparum* from the literature (Morgan *et al.*, 2005, Morgan *et al.*, 2004, Uthaipibull *et al.*, 2001) aligned against *P. yoelii* MSP1₁₉. This figure shows that residue 28 is conserved across the species. Uthaipibull *et al.* produced a Glu28→Ile variant which was found to affect inhibitory antibody binding (Uthaipibull *et al.*, 2001). This was in agreement with the *P. yoelii* data presented in this thesis for the Glu28→Lys and Glu28→Gln variant. The antibody binding studies imply that residue 28 is important for inhibitory antibody binding across the species and its conservation may be of functional importance. The NMR structure of the Glu28→Lys *P. yoelii* MSP1₁₉ variant presented in this thesis however suggests that residue 28 plays a vital structural role for *P. yoelii* MSP1₁₉. This could therefore suggest that residue 28 plays a vital structural role across the species because the variation to this residue in *P. falciparum* had a significant affect on antibody binding. The structure of Glu28→Ile from *P. falciparum* was not solved but it could be predicted from my data that it could have a significant structural affect on the protein. Residues 16 and 17 were found to be important in antibody binding in this thesis but variations in the equivalent residues by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) in *P. falciparum* did not show an important role. Residue 16 is not conserved across the species so this may explain the difference

for this residue. The residues shown to be important by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) for inhibitory antibody binding are mapped onto the NMR structure of *P. yoelii* MSP1₁₉ (shown in figure 10.7, panel B). This indicates that the residues identified by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) to be important for inhibitory antibody binding are located close to the residues shown to be important for antibody binding in this thesis.

The cross saturation NMR studies of Morgan *et al.* indicate that one of the interfaces for inhibitory antibody binding to *P. falciparum* MSP1₁₉ lies between residues 16 and 22 (shown in figure 10.6) (Morgan *et al.*, 2005, Morgan *et al.*, 2004). These data agree with the findings in this thesis which indicate that residues 16 and 17 are important for antibody binding. Residues 12 and 28 shown in this report to affect antibody binding lie outside the areas indicated by Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) to affect *P. falciparum* inhibitory antibody binding but there is a small number of residues located around these two residues in the NMR structure of *P. yoelii* MSP1₁₉ (shown in figure 10.7, panel C) that do affect inhibitory antibody binding. The epitope mapping NMR studies from Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) showing that residue 28 does not affect inhibitory antibody binding in *P. falciparum* but the surrounding residues do have an effect could support the structural NMR studies carried out in this thesis on *P. yoelii* MSP1₁₉ which suggested that the variation to residue 28 had caused significant structural perturbation and that this structural perturbation was responsible for the effect on antibody binding.

The data from Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) and Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) indicate that residues 12, 16, 17 and 28 shown in this thesis to affect *P. yoelii* MSP1₁₉ antibody binding lie within the same area of MSP1₁₉ as the residues identified as important for binding in *P. falciparum*. This suggests that there are specific areas that are important for inhibitory antibody binding across the species rather than the exact same residues being involved. This could imply that there is a common mechanism of action for the inhibitory antibodies across the species.

The identification of a common location for inhibitory antibody binding could be important in developing antigens for vaccination to specifically stimulate production of inhibitory antibodies. The potential identification of a common mechanism for

inhibitory antibody action could help direct studies to understand the mechanism of action for these antibodies and to develop new therapeutic strategies targeting MSP1₁₉.

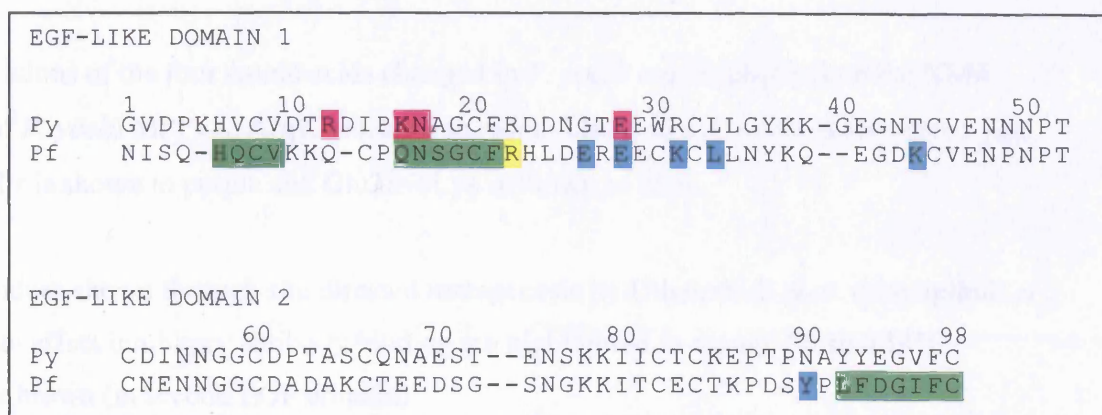


Figure 10.6: Alignment of *P. yoelii* (Py) and *P. falciparum* (Pf) MSP1₁₉ highlighting the four *P. yoelii* variants produced in this thesis relative to residues shown to affect *P. falciparum* MSP1₁₉ inhibitory antibody binding.

The residues altered in this thesis are highlighted in pink. The residues shown through site directed mutagenesis by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) to affect inhibitory antibody binding are highlighted in blue. The residues shown through cross saturation NMR studies to be in the binding site interface of inhibitory antibodies, 12.8 and 12.10 by Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) are highlighted in green. The residue highlighted in yellow was shown by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) and Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) to be important for inhibitory antibody binding. The alignment is based on figure 3 of Benjamin *et al.* (Benjamin *et al.*, 1999).

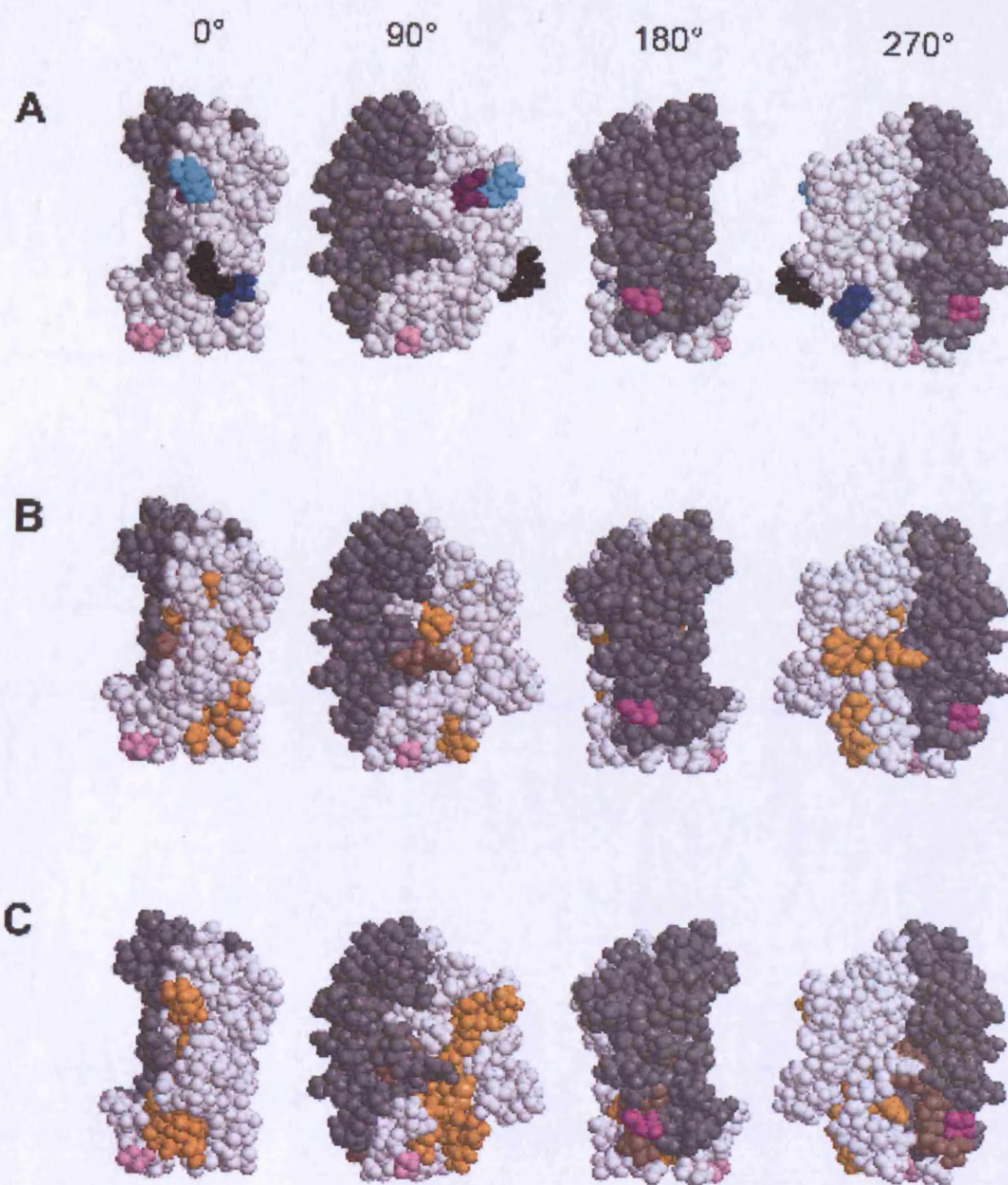
Figure 10.7: Important residues for inhibitory antibody binding in *P. falciparum* MSP1₁₉ mapped onto the NMR structure of *P. yoelii* MSP1₁₉.

A: The locations of the four amino acids changed in *P. yoelii* are highlighted on the NMR structure of *P. yoelii* MSP1₁₉: Arg12→Leu is shown in black, Lys16→Glu is shown in cyan, Asn17→His is shown in purple and Glu28→Lys is shown in blue.

B: The residues shown through site directed mutagenesis by Uthaipibull *et al.* (Uthaipibull *et al.*, 2001) to affect inhibitory antibody binding are highlighted in orange (in first EGF domain) or brown (in second EGF domain).

C: The residues shown through cross saturation NMR studies and chemical shift perturbation to be in the binding site interface of inhibitory antibodies 12.8 and 12.10 by Morgan *et al.* (Morgan *et al.*, 2005, Morgan *et al.*, 2004) are highlighted in orange (in first EGF domain) or brown (in second EGF domain).

The C-terminal residue is shown in bright pink and the N-terminal residue is shown in light pink.



10.5 **Overall implications of this work**

The data presented in this thesis could have implications for future research on MSP1₁₉ in *P. yoelii* and *P. falciparum* and for vaccine development involving protein engineering. In this thesis, I have shown that by altering a single residue (residue 28) on the surface of the protein, it can have a significant effect on the structure of the protein. This information is important for engineering proteins for vaccination because single amino acid changes are frequently made to proteins to help with large scale expression of the protein. For example, glycosylation sites and cleavage sites may be removed. This work has shown that a surface residue which had been predicted to be unlikely to have a large structural effect can affect the structure. The data shows that when making small changes to proteins for vaccine development it is vital to confirm that the structure remains intact even if the residue appears to be on surface of the protein. The data presented in this thesis for the residue 28 variant have also indicated this residue as a potential residue for conservation in MSP1₁₉ across the species for structural reasons and residue 28 is not part of the published conserved EGF structural motif. These data therefore suggest that there may be more structurally conserved residues in MSP1₁₉ and identifying these residues will help in determining what changes can be made to MSP1₁₉ to alter the immunogenicity for vaccination without altering the structure. It also suggests that you could have a residue like residue 28 that appears to be in the area of antibody binding that does not change under immune pressure because it plays a structurally important role.

The immunisation studies and ELISA antibody titres following immunisation with *P. yoelii* MSP1₁₉ presented in this thesis have shown that antibody titres in mice do not necessarily correlate with protection and that the fine specificity of the antibodies is more important in determining protection. This has implications for both immunisation studies in mice and the immune response to MSP1₁₉ in humans. These data suggest that measuring overall antibody titres will not provide an accurate determination of a protective immune response to malaria. It suggests that different indicators of protection need to be developed in order to test the effectiveness of engineered MSP1₁₉ proteins as vaccines.

The comparison of the antibody binding sites in *P. falciparum* and *P. yoelii* MSP1₁₉ has suggested a common area for antibody binding across the species. This could help direct further research with *P. falciparum* MSP1₁₉ which could help in the development of antigens for vaccination to specifically stimulate the production of inhibitory antibodies. The data could also help to target studies to understand the mechanism of action of the antibodies to help develop new therapeutic strategies targeting MSP1₁₉.

10.6 Future Work

The data presented in this thesis have identified important residues for antibody binding and a potentially important structural residue and have identified areas of the protein that are important in the ability of MSP1₁₉ to provide protection against parasite challenge. The data have suggested conservation across the species of important areas for antibody binding to the protein. The data have raised questions and indicated future research directions to further understanding of antibody binding to MSP1₁₉ across the species.

The immunisation studies presented have shown that the fine specificity of the antibody response to MSP1₁₉ is important in protection against parasite challenge and that immunisation with residue 28 MSP1₁₉ variant proteins does not protect against challenge infection. In order to confirm the theory that the important antibodies for protection are raised towards the areas of the protein that are altered in the residue 28 variants passive immunisation studies could be carried out. This would involve passive immunisation of mice with the antibodies generated to the residue 28 MSP1₁₉ variant proteins and then challenging the mice with *P. yoelii* YM parasites. If all the antibodies that are involved in protection against the parasite are made to the area of the protein altered in the residue 28 variant you may expect that the antibodies would not be able to bind to the native MSP1₁₉ on the parasite and therefore not give a protective immune response.

The 2D ¹⁵N-HSQC data for the Glu28→Gln MSP1₁₉ variant suggested that it caused less structural perturbation than the Glu28→Lys MSP1₁₉ variant. Further structural NMR studies could be carried out to fully understand the structural changes caused by the Glu28→Gln variation. This could help to explain why Glu28→Gln MSP1₁₉ variation causes significant structural perturbation whereas the Arg12→Leu variation which is located in the same area of the protein does not significantly alter the structure.

Transfection studies of the *P. yoelii* parasite could be carried out to introduce the residue 28 variations to examine whether the variations have an effect on the viability of

the parasite *in vivo*. If the parasite is viable immunisation studies immunising the mice with the Glu28→Lys MSP1₁₉ variant protein and challenging with the residue 28 variant parasite could be performed to see if the protein is able to protect against homologous parasite challenge.

In this thesis I have focused on antibodies that protect against parasite challenge with *P. yoelii* YM and compared the results to *P. falciparum* inhibitory antibodies. Assays need to be developed to identify if there are inhibitory, blocking and neutral antibodies against *P. yoelii* infection like those that occur in humans following *P. falciparum* malaria infection. If blocking antibodies were identified, the site directed mutagenesis and antibody binding study approach that was taken in this thesis could be applied to mapping the epitopes for blocking antibodies. Immunisation studies could be carried out to determine if it is possible to knock out the binding sites for blocking antibodies while maintaining the ability of MSP1₁₉ to protect against parasite challenge.

In order to confirm whether the data gained in this thesis for *P. yoelii* MSP1₁₉ can be translated to *P. falciparum* the variations to the equivalent residues could be made to *P. falciparum* MSP1₁₉ to see if they affect binding to the inhibitory antibodies. Residue 28 (residue 26 in *P. falciparum*) was identified for study in this thesis because it was a important residue for *P. falciparum* inhibitory antibody binding in the site directed mutagenesis studies of Uthaipibull *et al.* (Uthaipibull *et al.*, 2001). However the data I have presented in this thesis have suggested that in *P. yoelii* residue 28 has a vital structural role therefore the role of the equivalent residue on the structure of *P. falciparum* should be investigated to examine whether this residue has a conserved structural function across the species. This could involve making a *P. falciparum* Glu26→Lys MSP1₁₉ variant and carrying out 2D ¹⁵N-HSQC NMR analysis to look for any potential structural perturbation and if there was significant structural perturbation, 3D NMR analysis could be carried out to solve the structure of the variant and identify the structural role of the residue.

Acknowledgements

I would like to thank the following people for their help and support with this project. My supervisor Tony Holder for being my supervisor, providing support, encouragement and useful suggestions throughout my PhD and during my thesis preparation. Jean Langhorne for being my second supervisor and providing advice. Madhu Kadekoppala and Sola Ogun for being my practical supervisors and for their help with the immunisation studies including carrying out the animal handling. Irene Ling for providing BL21 cells containing the pGEX3X wildtype MSP1₁₉. Brian YimLim for providing BL21 gold cells containing pGEX3X vector and for advice on the expression of GST proteins. Lilian Spencer for making the B6, F5 and B10 antibodies. Chariat Uthaipibull for advice with site directed mutagenesis and BIAcore. Jason Billington in Virology for teaching me how to use the BIAcore machine and providing help and support. The Division of Parasitology and the Biomolecular NMR Centre for providing helpful suggestions and advice for experiments. Berry Birdsall for her help with running and setting up NMR experiments, help with solving the NMR structure, help with analysing NMR data and for teaching me how to analyse NMR data, solve a protein structure by NMR and use Linux. I would also like to thank Berry for her continued support throughout the NMR structural calculations, enthusiasm during the difficult times when the structures didn't look like they would converge and for advice regarding the preparation of the NMR chapters in this thesis. Geoff Kelly for help with setting up NMR experiments and NMR software. Tom Frenkiel and Alain Oregioni for providing pulse sequences for the NMR experiments and for their help and advice. Ceseri de Chiara and Flor Garcia for answering ARIA questions. John McCornmick for labelled protein preparation advice and help with freeze-drying.

I would also like to thank my family and partner for their continuous support and encouragement throughout my PhD and the preparation of this thesis.



Bibliography

- Al-Yaman, F., Genton, B., Kramer, K. J., Chang, S. P., Hui, G. S., Baisor, M. & Alpers, M. P. (1996). Assessment of the role of naturally acquired antibody levels to *Plasmodium falciparum* merozoite surface protein-1 in protecting Papua New Guinean children from malaria morbidity. *Am. J. Trop. Med. Hyg.* **54**, 443-448.
- APPMG (2005). Tackle Malaria Today, Give tomorrow a chance. London: House of Commons All-Party Parliamentary Malaria Group.
- Babon, J., Morgan, W. D., Kelly, G., Eccleston, J. F., Feeney, J. & Holder, A. A. (2007). Structural studies on *Plasmodium vivax* merozoite surface protein-1. *Molecular and Biochemical Parasitology* **153**, 31-40.
- Bartels, C., Xia, T., Billiter, N., Guntert, P. & Wuthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **6**, 1-10.
- Benjamin, P. A., Ling, I. T., Clotey, G., Spencer Valero, L. M., Fleck, S. L., Walliker, D., Morgan, W. D., Birdsall, B., Feeney, J. & Holder, A. A. (1999). Antigenic and sequence diversity at the C-terminus of the merozoite surface protein-1 from the rodent malaria isolates, and the binding of protective monoclonal antibodies. *Molecular and Biochemical Parasitology* **104**, 147-156.
- Bentley, G. A. (2006). Functional and immunological insights from the three-dimensional structures of *Plasmodium* surface proteins. *Current Opinion in Microbiology* **9**, 1-6.
- Birnboim, H. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513-23.
- Blackman, M. J., Heidrich, H. G., Donachie, S., McBride, J. S. & Holder, A. A. (1990). A Single Fragment of a Malaria Merozoite Surface Protein Remains on the Parasite During Red Cell Invasion and is the Target of Invasion Inhibiting Antibodies. *Journal of Experimental Medicine* **179**, 379-382.
- Blackman, M. J. & Holder, A. A. (1992). Secondary processing of the *Plasmodium falciparum* Merozoite Surface Protein-1 (MSP1) by a calcium-dependent membrane-bound serine protease-shedding of MSP1₃₃ as a noncovalently associated complex with other fragments of the MSP1. *Mol. Biochem. Parasitol.* **50**, 307-316.
- Blackman, M. J., Scott-Finnigan, T. J., Shai, S. & Holder, A. A. (1994). Antibodies Inhibit the Protease-mediated Processing of a Malaria Merozoite Surface Protein. *J. Exp. Med.* **180**, 389-393.
- Blackman, M. J., Whittle, H. & Holder, A. A. (1991). Processing of the *Plasmodium falciparum* major merozoite surface protein-1: identification of a 33-kilodalton

secondary processing product which is shed prior to erythrocyte invasion. *Molecular and Biochemical Parasitology* **49**, 35-44.

- Branch, O. H., Udhayakumar, V., Hightower, A. W., Oloo, A. J., Hawley, W. A., Nahlen, B. L., Bloland, P. B., Kaslow, D. C. & Lal, A. A. (1998). A longitudinal investigation of IgG and IgM antibody responses to the merozoite surface protein-1 19-kilodalton domain of *Plasmodium falciparum* in pregnant women and infants: associations with febrile illness, parasitemia, and anemia. *Am. J. Trop. Med. Hyg.* **58**, 211-219.
- Breman, J. G. (2001). The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. *Am. J. Trop. Med. Hyg.* **64**, 1-11.
- Breman, J. G., Egan, A. & Keusch, G. T. (2001). The intolerable burden of malaria: a new look at the numbers. *Am. J. Trop. Med. Hyg.* **64**, iv-vii.
- Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M. & Skelton, N. J. (2007). Protein NMR spectroscopy : principles and practice 2edn. Amsterdam: Elsevier Academic Press.
- Chiang, P. K., Bujnicki, J. M., Su, X. & Lanar, D. E. (2006). Malaria: Therapy, Genes and Vaccines. *Current Molecular Medicine* **6**, 309-326.
- Chitarra, V., Holm, I., Bentley, G. A., Petres, S. & Longacre, S. (1999). The crystal structure of C-terminal merozoite surface protein 1 at 1.8Å resolution, a highly protective malaria vaccine candidate. *Mol. Cell* **3**, 457-464.
- Clare, J. J., Rayment, F., Ballantine, S., Sreekrishna, K. & Romanos, M. A. (1991). Production of mouse epidermal growth factor in yeast: high level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* **105**, 205-212.
- Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289-302.
- Creighton, T. E. (1997). Proteins: Structures and Molecular Properties, 2 edn. New York: W.H. Freeman and Company.
- Daly, T. M. & Long, C. A. (1993). A Recombinant 15-Kilodalton Carboxyl-Terminal Fragment of *Plasmodium yoelii yoelii* 17XL Merozoite Surface Protein 1 Induces a Protective Immune Response in Mice. *Infection and Immunity* **61**, 2462-2467.
- Dayringer, H., Tramontano, A., Sprang, S. & Fletterick, R. (1986). Insight an interactive molecular graphics package. *J. Mol. Graphics* **4**, 82-87.
- de Koning-Ward, T. F., O'Donnell, R. A., Drew, R. D., Thomson, R., Speed, T. P. & Crabb, B. S. (2003). A New Rodent Model to Assess Blood Stage Immunity to the Plasmodium falciparum Antigen Merozoite Surface Protein 1₁₉ Reveals a Protective Role for Invasion Inhibitory Antibodies. *J. Exp. Med.* **198**, 869-875.

- Dekker, C., Uthaipibull, C., Calder, L. J., Lock, M. J., Grainger, M., Morgan, W. D., Dodson, G. G. & Holder, A. A. (2004). Inhibitory and neutral antibodies to *Plasmodium falciparum* MSP1₁₉ form ring structures with their antigen. *Molecular and Biochemical Parasitology* **137**, 143-149.
- Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293.
- Dodoo, D., Theander, T. G., Kurtzhals, A. L., Koram, K., Riley, E. M., Akanmori, B. D., Nkrumah, F. K. & Hviid, L. (1999). Levels of Antibody to Conserved Parts of *Plasmodium falciparum* Merozoite Surface Protein 1 in Ghanaian Children Are Not Associated with Protection from Clinical Malaria. *Infection and Immunity* **67**, 2131-2137.
- Edwards, A. J. & Reid, D. (2000). Introduction to NMR of proteins. *Current Protocols in Protein Science* **Unit 17.5**, 17.5.1-17.5.39.
- Egan, A. F., Morris, J., Barnish, G., Allen, S., Greenwood, B. M., Kaslow, D. C., Holder, A. A. & Riley, E. M. (1996). Clinical Immunity to *Plasmodium falciparum* Malaria Is Associated with Serum Antibodies to the 19-kDa C-terminal Fragment of the Merozoite Surface Antigen, PfMSP-1. *The Journal of Infectious Diseases* **173**, 765-769.
- Fernandez, C. & Wider, G. (2003). TROSY in NMR studies of the structure and function of large biological molecules. *Current Opinion in Structural Biology* **13**, 570-580.
- Freifelder, D. (1999). *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2 edn. New York: W.H Freeman and company.
- Garman, S. C., Simcoke, W. N., Stowers, A. W. & Garboczi, D. N. (2003). Structure of the C-terminal domains of merozoite surface protein-1 from plasmodium knowlesi reveals a novel histidine binding site. *J. Biol. Chem.* **278**, 7264-7269.
- Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
- Goddard, T. D. & Kneller, D. G. SPARKY 3, university of California, San Francisco.
- Guevara Patino, J. A., Holder, A. A., McBride, J. S. & Blackman, M. J. (1997). Antibodies that inhibit Malaria Merozoite Surface Protein-1 Processing and Erythrocyte Invasion Are Blocked by Naturally Acquired Human Antibodies. *J. Exp. Med.* **186**, 1689-1699.
- Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb viewer: An environment for comparative protein modelling. *Electrophoresis* **18**, 2714-2723.

- Hoffman, S. L. & Miller, L. H. (1996). Perspectives on Malaria Vaccine Development. In *Malaria Vaccine Development: A Multi-Immune Response Approach*, pp. 1-13. Edited by S. L. Hoffman. Washington, D.C.: ASM Press.
- Hogh, B., Marbiah, N. T., Burghaus, P. A. & Andresen, P. K. (1995). Relationship between maternally derived anti-*Plasmodium falciparum* antibodies and risk of infection and disease in infants living in an area of Liberia, west Africa, in which is highly endemic. *Infection and Immunity* **63**.
- Holder, A. A. & Blackman, M. J. (1994). What is the function of MSP-1 on the Malaria Merozoite? *Parasitology Today* **10**, 182-184.
- Holder, A. A., Chappel, J. A. & Blackman, M. J. (1994). Malaria Merozoite Surface Protein-1 Structure and Processing: Targets to interfere with the parasite life cycle. In *Recombinant and Synthetic Vaccines*, pp. 99-106. Edited by G. P. Talwar, K. V. S. Rao & V. S. Chaunhan. New Delhi: Narosa Publishing House.
- Koradi, R., Billeter, M. & Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51-55.
- Laskowski, R., Rullmann, J., MacArthur, M., Kaptein, R. & Thornton, J. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486.
- Ling, I. T., Ogun, S. A. & Holder, A. A. (1994). Immunization against malaria with a recombinant protein. *Parasite immunology* **16**, 63-67.
- Ling, I. T., Ogun, S. A., Momin, P., Richards, R. L., Garcon, N., Cohen, J., Ballou, W. R. & Holder, A. A. (1997). Immunisation against the murine malaria parasite *Plasmodium yoelii* using a recombinant protein with adjuvants developed for clinical use. *Vaccine* **15**, 1562-1567.
- Linge, J. P. & Nilges, M. (1999). Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. *J. Biomol. NMR* **13**, 51-59.
- Linge, J. P., O' Donoghue, S. & Nilges, M. (2001). Assigning Ambiguous NOEs with ARIA *Methods in Enzymology* **339**, 71-90.
- McBride, J. S. & Heidrich, H. G. (1987). Fragments of the polymorphic Mr 185000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites form an antigenic complex. *Mol. Biochem. Parasitol.* **23**, 71-84.
- Miller, L. H., Howard, R. J., Carter, R., Good, M. F., Nussenzweig, V. & Nussenzweig, R. S. (1986). Research Toward Malaria Vaccines. *Science* **234**, 1349-1356.
- Morgan, W. D., Birdsall, B., Frenkiel, T. A., Gradwell, M. G., Burghaus, P. A., Syed, S. E. H., Uthaipibull, C., Holder, A. A. & Feeney, J. (1999). Solution Structure of an EGF Module Pair from the *Plasmodium falciparum* Merozoite Surface Protein 1. *J. Mol. Biol* **289**, 113-122.

- Morgan, W. D., Frenkiel, T. A., Lock, M. J., Grainger, M. & Holder, A. A. (2005). Precise Epitope Mapping of Malaria Parasite Inhibitory Antibodies by TROSY NMR Cross-Saturation. *Biochemistry* **44**, 518-523.
- Morgan, W. D., Lock, M. J., Frenkiel, T. A., Grainger, M. & Holder, A. A. (2004). Malaria Parasite-inhibitory antibody epitopes in *Plasmodium falciparum* merozoite surface protein-1₁₉ mapped by TROSY NMR. *Molecular and Biochemical Parasitology* **138**, 29-36.
- Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol* **245**, 645-660.
- Nilges, M., Macias, M. J., O' Donoghue, S. & Oschkinat, H. (1997). Automated NOESY Interpretation with Ambiguous Distance Restraints: The Refined NMR Solution Structure of the Pleckstrin Homology Domain from β -spectrin. *J. Mol. Biol* **269**, 408-422.
- Nilges, M. & O' Donoghue, S. (1998). Ambiguous NOEs and automated NOE assignment. *Prog. NMR Spect.* **32**, 107-139.
- Norwood, T. J., Boyd, J. & Campbell, I. D. (1989). Improved resolution in ^1H - ^{15}N correlation experiments. *FEBS* **255**, 369-371.
- Nwuba, R., Sodeinde, O., Anumudu, C. I., Omosun, Y. O., Odaibo, A. B., Holder, A. A. & Nwagwu, M. (2002). The Human Immune Response to *Plasmodium falciparum* includes both Antibodies that Inhibit Merozoite Surface Protein 1 Secondary Processing and Blocking Antibodies. *Infection and Immunity* **70**, 5328-5331.
- Outchkourov, N. S., Steikema, W. J. & Jongsma, M. A. (2002). Optimization of the expression of Equistatin in *Pichia pastoris*. *Protein Expression and Purification* **24**, 18-24.
- Peitsch, M. C. (1995). Protein modelling by E-mail. *Bio/Technology* **13**, 658-660.
- Piotto, M., Saudek, V. & Sklenár, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661-665.
- Pizarro, J. C., Chitarra, V., Verger, D., Holm, I., Petres, S., Dartville, S., Nato, F., Longacre, S. & Bentley, G. A. (2003). Crystal Structure of a Fab Complex Formed with Pfmsp1-19, the C-terminal Fragment of Merozoite Surface Protein 1 from *Plasmodium Falciparum*: A Malaria Vaccine Candidate. *J. Mol. Biol* **328**, 1091.
- Riley, E. M., Allen, S. L., Wheeler, J. G., Blackman, M. J., Bennett, S., Takacs, B., Schonfeld, H. J., Holder, A. A. & Greenwood, B. M. (1992). Naturally acquired cellular and humoral immune responses to the major merozoite surface antigen (Pf MSP1) of *Plasmodium falciparum* are associated with reduced malaria morbidity. *Parasite immunology* **14**, 321-337.

- Riley, E. M., Morris-Jones, S., Blackman, M. J., Greenwood, B. M. & Holder, A. A. (1993). A longitudinal study of naturally acquired cellular and humoral immune responses to a merozoite surface protein (MSP1) of *Plasmodium falciparum* in an area of seasonal malarai transmission. *Parasite immunology* **15**, 513-524.
- Romanos, M. A., Scorer, C. A. & Clare, J. J. (1992). Foreign Gene Expression in Yeast: a review. *Yeast* **8**, 423-488.
- Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modelling server. *Nucleic Acids Research* **31**, 3381-3385.
- Spencer Valero, L. M., Ogun, S. A., Fleck, S. L., Ling, I. T., Scott-Finnigan, T. J., Blackman, M. J. & Holder, A. A. (1998). Passive Immunization with Antibodies against Three Distinct Epitopes of *Plasmodium yoelii* Merozoite Surface Protein 1 Suppresses Parasitemia. *Infection and Immunity* **66**, 3925-3930.
- Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J., Smith, P. L., Wierschke, J. D., Subramaniam, A. & Birkenberger, L. A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190**, 55-62.
- Uthaipibull, C., Aufiero, B., Syed, S. E. H., Hansen, B., Guevara Patino, J. A., Angov, E., Ling, I. T., Fegeding, K., Morgan, W. D., Ockenhouse, C., Birdsall, B., Feeney, J., Lyon, J. A. & Holder, A. A. (2001). Inhibitory and Blocking Monoclonal Antibody Epitopes on Merozoite Surface Protein 1 of the Malaria Parasite *Plasmodium falciparum*. *Journal of Molecular Biology* **307**, 1381-1394.
- Valadon, P. (2007). RasTop 2.2.
- Wan Omar, A., Roslaini, A. M., Ngah, Z. U., Azahari, A. A., Zahedi, M. & Baharudin, O. (2007). A recombinant 19kDa *Plasmodium berghei* merozoite surface protein 1 formulated with alum induces protective immune response in mice. *Tropical Biomedicine* **24**, 119-126.
- Woo, J. H., Liu, Y. Y., Mathias, A., Scott, S., Wang, Z., Thompson, J. & Neville, D. M. (2002). Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in *Pichia pastoris*. *Protein Expression and Purification* **25**, 270-282.
- Wouters, M. A., Rigoutsos, I., Chu, C. K., Feng, L. L., Sparrow, D. B. & Dunwoodie, S. L. (2005). Evolution of distinct EGF domains with specific functions. *Protein Science* **14**, 1091-1103.
- Zhang, Z. H., Jiang, P. H., Li, N. J., Shi, M. & Huang, W. (2005). Oral vaccination of mice against rodent malaria with recombinant *Lactococcus lactis* expressing MSP-1₁₉. *World Journal of Gastroenterology* **11**, 6975-6980.

